

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 06:56:19 ; Search time 3876.86 Seconds
(without alignments)
10990.348 Million cell updates/sec

Title: US-10-009-966C-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tggtcggaattaccgacca.....atgacttcattctgattctc 901

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	899	99.8	3294	6	AX068012 Sequence
2	899	99.8	4135	6	AX068019 Sequence
3	899	99.8	4312	6	AX068013 Sequence
4	893.8	98.1	3293	8	AF376772S1 Nicotiana
5	76.4	8.5	1705	8	NP056861 Nicotiana p
6	76.4	8.5	7299	8	AB012638 Nicotiana
7	72.4	8.0	3293	8	AF376772S1 Nicotiana
8	72.4	8.0	3294	6	AX068012 Sequence
9	72.4	8.0	4135	6	AX068019 Sequence
10	72.4	8.0	4312	6	AX068013 Sequence
11	71.8	8.0	159636	8	AC090485 Genomic S
12	69.6	7.7	2997	8	AF233297 Nicotiana
13	69.2	7.7	120480	8	AF006654 Lotus cor
14	68.8	7.5	7869	8	LLVFNAG L. luteus tr
15	68	7.5	628	8	NAU40606 Nicotiana a
16	67.2	7.5	132473	8	AC136451 Medicago
17	67.2	7.5	125460	8	AP004850 Oryza sat
18	67.2	7.5	176446	8	AP005514 Oryza sat
19	66.4	7.4	142196	8	AC130728 Oryza sat

c	20	66	7.3	20264	8	AC079675	Arabidops
c	21	65.8	7.3	83796	8	AC004165	Arabidops
c	22	65.6	7.3	135334	8	AC146554	Medicago
c	23	65.2	7.2	125800	8	AC118670	Genomic s
c	24	65.2	7.2	139714	8	AC137931	Oryza sat
c	25	65	7.2	100773	8	AF466201	Ar466201 Sorghum b
c	26	64.8	7.2	226	8	AY019961	AY019961 Oryza sat
c	27	64.8	7.2	1140	6	CQ755596	Sequence
c	28	64.8	7.2	1140	6	CQ755618	Sequence
c	29	64.8	7.2	1140	6	CQ757793	Sequence
c	30	64.8	7.2	1140	6	CQ757815	Sequence
c	31	64.8	7.2	1140	6	AX764627	Sequence
c	32	64.8	7.2	1140	6	AX764649	Sequence
c	33	64.8	7.2	82411	8	AC068809	U93215 Arabidops S
c	34	64.8	7.2	104861	8	U93215	U93215 Arabidops S
c	35	64.4	7.1	112689	2	AC149818	Ze mays
c	36	64.4	7.1	136471	8	AP003622	AP003622 Oryza sat
c	37	64.4	7.1	175153	8	AP003977	AP003977 Oryza sat
c	38	64	7.1	160	8	SSTRNG	X06795 Sorghum bic
c	39	64	7.1	5945	6	AX344987	AX344987 Sequence
c	40	64	7.1	139155	2	AC149815	AC149815 Zea mays
c	41	64	7.1	266544	3	AC116956	AC116956 Dictyoste
c	42	63.6	7.1	120787	8	AC002292	AC002292 Arabidops
c	43	63.6	7.1	124214	2	AC130803	AC130803 Medicago
c	44	63.4	7.0	399	11	EX545998	EX545998 Arabidops
c	45	63.4	7.0	458	11	AL823587	AL823587 Arabidops

ALIGNMENTS

RESULT 1
AX068012
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Nicotiana tabacum (common tobacco)

AX068012
Sequence 1 from Patent WO0077187.
AX068012
AX068012.1 GI:12329820

3294 bp DNA linear PAT 19-JAN-2001

REFERENCE
AUTHORS
TITLE
JOURNAL

Roitsch, T.D.
Promoter system and production and use of the same
Patent: WO 0077187-A 1 21-DEC-2000;
Roitsch, Thomas, D. (1998)

FEATURES
source

Location/Qualifiers
1..3294
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN

Query Match 99.8%; Score 899; DB 6; Length 3294;
Best Local Similarity 100.0%; Pred. No. 3.2e-154; Indels 0; Gaps 0;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGGTCGGAATTAACCGACCAACTTTGGTCGTCATTAATAATCAAAAATATTTGTTAAA 60
Db 2255 TGGTCGGAATTAACCGACCAACTTTGGTCGTCATTAATAATCAAAAATATTTGTTAAA 2314
QY 61 AAAAAACGACCAAGTTGTCGGTATTTAATTAATTAATAAAGAAATTCACATCTCGG 120
Db 2315 AAAAAACGACCAAGTTGTCGGTATTTAATTAATTAATAAAGAAATTCACATCTCGG 2374
QY 121 GAATCGAACCGGGGTCTGTACTATGCGAAGTACTATTTCTACCACTAGACCATTTGTTCA 180
Db 2375 GAATCGAACCGGGGTCTGTACTATGCGAAGTACTATTTCTACCACTAGACCATTTGTTCA 2434
QY 181 TTTTGTGTTTAAAGCTGCTTTTATTTGTTTATATCTTTTATATATTTTGTGACGAAA 240

parent
PG

Db 2435 TTTGTTTAAAGACTGCTTTTATTGATTATATCTTTAAATATATATTTTTCACGAAA 2494

Qy 241 ATACCGACCAAGTTCGTGATTTTATTAAGAAGTAATTAATTAACCAAGTTCGTG 300

Db 2495 ATACCGACCAAGTTCGTGATTTTATTAAGAAGTAATTAATTAACCAAGTTCGTG 2554

Qy 301 ATTTTAAATGATCCGCGAATTAACCGACCAATTTTGGTAGTTTTCGTAATTA 360

Db 2555 ATTTTAAATGATCCGCGAATTAACCGACCAATTTTGGTAGTTTTCGTAATTA 2614

Qy 361 TTTTATTTTAAATGATCCGCGAATTAACCGACCAATTTTGGTAGTTTTCGTAATTA 420

Db 2615 TTTTATTTTAAATGATCCGCGAATTAACCGACCAATTTTGGTAGTTTTCGTAATTA 2674

Qy 421 TCGCGGACCAAAATAGTTTCCGCGATTTTTCGCGAAGTAAAGAAACCGACCAAGTTG 480

Db 2675 TCGCGGACCAAAATAGTTTCCGCGATTTTTCGCGAAGTAAAGAAACCGACCAAGTTG 2734

Qy 481 TCGGTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 540

Db 2735 TCGGTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2794

Qy 541 GTCGTTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600

Db 2795 GTCGTTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2854

Qy 601 TCGCGAATTTCTAGTAGTACGCAACCTTGAAGCTTCGGGAGAAATTTTGTATATGA 660

Db 2855 TCGCGAATTTCTAGTAGTACGCAACCTTGAAGCTTCGGGAGAAATTTTGTATATGA 2914

Qy 661 TATGTATATCTTAAATGATTTTAAAGAAACCGACCAATTTTGTATATGA 720

Db 2915 TATGTATATCTTAAATGATTTTAAAGAAACCGACCAATTTTGTATATGA 2974

Qy 721 TTTAGGGGCACTAGATGAGCAGAAATACGTTTCTCGTCGCTAAAGAAATTTTGTATATGA 780

Db 2975 TTTAGGGGCACTAGATGAGCAGAAATACGTTTCTCGTCGCTAAAGAAATTTTGTATATGA 3034

Qy 781 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

Db 3035 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3094

Qy 841 TAAACTATAGACTCGCTTTTATAGCACTTTTAAAGAACTATGACTTCTGATTTCT 900

Db 3095 TAAACTATAGACTCGCTTTTATAGCACTTTTAAAGAACTATGACTTCTGATTTCT 3154

RESULT 2

AX068019

LOCUS AX068019 4135 bp DNA linear PAT 19-JAN-2001

DEFINITION Sequence 8 from Patent WO007187.

ACCESSION AX068019

VERSION AX068019.1 GI:12329827

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Roitsch, T.D.

TITLE Promoter system and production and use of the same

JOURNAL Patent: WO 007187-A. 8 21-DEC-2000.

Roitsch, Thomas, Dr. (DE)

FEATURES

source

1. .4135

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Fusion aus Promotor und codierender Sequenz in antisense-Orientierung"

ORIGIN

Query Match 99.8%; Score 899; DB 6; Length 4135;

Best Local Similarity 100.0%; Pred. No. 3.1e-154;

Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGTCGGGAAATTTACCGACCAACTTTGGTCGGTCAATTAATAATCAAAAAAATATTGTAAA 60

Db 2255 TGGTCGGGAAATTTACCGACCAACTTTGGTCGGTCAATTAATAATCAAAAAAATATTGTAAA 2314

Qy 51 AAAAAACCGACCAAGTTGATCGGTATTTTAATTAATTAATAATAATAATAATAATAATA 120

Db 2315 AAAAAACCGACCAAGTTGATCGGTATTTTAATTAATTAATAATAATAATAATAATAATA 2374

Qy 121 GAATCGAAACCGGGGCTCTGATCTATGCGCAAGATATCTATTCTACCACTAGACCATTTGTTCA 180

Db 2375 GAATCGAAACCGGGGCTCTGATCTATGCGCAAGATATCTATTCTACCACTAGACCATTTGTTCA 2434

Qy 181 TTTGTTTAAAGACTGCTCTTTTATTTGATTATATCTCTTTAAATTAATTTTTCACGAAA 240

Db 2435 TTTGTTTAAAGACTGCTCTTTTATTTGATTATATCTCTTTAAATTAATTTTTCACGAAA 2494

Qy 241 ATACCGACCAAGTTCGTGATTTTATTAAGAAGTAATTAATTAATTAATTAATTAATTAATTA 300

Db 2495 ATACCGACCAAGTTCGTGATTTTATTAAGAAGTAATTAATTAATTAATTAATTAATTAATTA 2554

Qy 301 ATTTTAAATGATCCGCGAATTTAAACCGACCAATTTTGGTAGTTTTCGTAATTAATA 360

Db 2555 ATTTTAAATGATCCGCGAATTTAAACCGACCAATTTTGGTAGTTTTCGTAATTAATA 2614

Qy 361 TTTTATTTTAAATGATCCGCGAATTTAAACCGACCAATTTTGGTAGTTTTCGTAATTAATA 420

Db 2615 TTTTATTTTAAATGATCCGCGAATTTAAACCGACCAATTTTGGTAGTTTTCGTAATTAATA 2574

Qy 421 TCGCGGACTCAAAAATAGTTTCCGCGATTTTTCGCGCAAGAAACCGACCAAGTTG 480

Db 2675 TCGCGGACTCAAAAATAGTTTCCGCGATTTTTCGCGCAAGAAACCGACCAAGTTG 2734

Qy 481 TCGGTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 540

Db 2735 TCGGTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2794

Qy 541 GTCGTTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600

Db 2795 GTCGTTTTCGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2854

Qy 601 TCGCGAATTTCTAGTAGTACGCAACCTTGAAGCTTCGGGAGAAATTTTGTATATGA 660

Db 2855 TCGCGAATTTCTAGTAGTACGCAACCTTGAAGCTTCGGGAGAAATTTTGTATATGA 2914

Qy 661 TATGTATATCTTAAATGATTTTAAAGAAACCGACCAATTTTGTATATGA 720

Db 2915 TATGTATATCTTAAATGATTTTAAAGAAACCGACCAATTTTGTATATGA 2974

Qy 721 TTTAGGGGCACTAGATGAGCAGAAATACGTTTCTCGTCGCTAAAGAAATTTTGTATATGA 780

Db 2975 TTTAGGGGCACTAGATGAGCAGAAATACGTTTCTCGTCGCTAAAGAAATTTTGTATATGA 3034

Qy 781 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

Db 3035 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3094

Qy 841 TAAACTATAGACTCGCTTTTATAGCACTTTTAAAGAACTATGACTTCTGATTTCT 900

Db 3095 TAAACTATAGACTCGCTTTTATAGCACTTTTAAAGAACTATGACTTCTGATTTCT 3154

Qy 901 C 901

Db 3155 C 3155

RESULT 3

AX068013

parent pct


```

/mol_type="genomic DNA"
/db_xref="taxon:4096"
/clone="lambda NSCbb1"
/tissue_type="leaf"
/clone_in="lambda DASH genomic library"
complement(1870..2673)
/gene="Lhcb1*6"
complement(1870..2673)
/gene="Lhcb1*6"
/function="antenna complex"
/codon_start=1
/product="light harvesting chlorophyll a/b-binding protein"
/protein_id="BAA25392.1"
/db_xref="GI:3036948"
/translations="MAASTMALSSSFAGKAVKLSPPSSSEITGNGKVTMRKTSKAKP
VSSGWPYGDPRVKYLPFSGESPSYLTGFPGDYGMDTAGLSADPETFAKNELEVI
HCRWALGCGVPELLARNGVKFGEAVWFAGSOFSEGGDLIGNPSLVAQSILA
AIWACQVLMGAVEGYRVAGGPLGEVVDPLYPGGSFDP/LGLAEDPEAFALVKYEIKN
GLRMSFMFGFVQALVTGKPLENLADHLADPVNNNAWAYATNFVPGK"
complement(2572..2673)
/gene="Lhcb1*6"
/notes="putative"
complement(1873..2571)
/gene="Lhcb1*6"
/product="light harvesting chlorophyll a/b-binding protein"
/notes="putative"
complement(2760..2767)
/notes="Lhcb1*6
putative"
4272..4278
/notes="Lhcb1*5
putative"
4363..5163
/gene="Lhcb1*5"
4363..5163
/gene="Lhcb1*5"
/function="antenna complex"
/codon_start=1
/product="light harvesting chlorophyll a/b-binding protein"
/protein_id="BAA25393.1"
/db_xref="GI:3036949"
/translations="MAASTMALSSSFAGKAVKLSPPSSSEITGNGKVTMRKTSKAKPV
SSGSLYGPBRVYLPFSGESPSYLTGFPGDYGMDTAGLSADPETFAKNELEVIH
CRWALGCGVPELLARNGVKFGEAVWFAGSOFSEGGDLIGNPSLVAQSILA
IWACQVLMGAVEGYRVAGGPLGEVVDPLYPGGSFDP/LGLAEDPEAFALVKYEIKN
GLRMSFMFGFVQALVTGKPLENLADHLADPVNNNAWAYATNFVPGK"
4363..4461
/gene="Lhcb1*5"
/notes="putative"
4462..5160
/gene="Lhcb1*5"
/product="light harvesting chlorophyll a/b-binding protein"
/notes="putative"

```

ORIGIN

```

Query Match      8.5%; Score 76.4; DB 8; Length 7299;
Best Local Similarity 79.5%; Pred. No. 0.00024;
Matches 89; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 628 CCTGTAGCTCGGGAGAAATTTGTATGTATGTATGTATGTATCTTAAATGATTAT 687
    |||||
Db 3928 CCCTAACTTCAGCAAACTCTGTATGTATGTATGTATGTATCTTAAATGATTAT 3869
    |||||

QY 688 TTAAGAACGNGCACCTCGAATACTAGAACCTTTAGGGCGCACTAGATGAG 739
    |||||
Db 3868 TTAATCACTTGACACCTTAATGCTAAAGCCTTTAGGGCGCAGCATCAG 3817
    |||||

```

RESULT 7

```

AF376772S1/c AF376772S1 3293 bp DNA linear PLN 29-MAY-2001
LOCUS Nicotiana tabacum extracellular invertase Nin8 (Nin8) gene,
DEFINITION promoter and partial cds.
ACCESSION AF376772
VERSION AF376772.1 GI:14211752
KEYWORDS 1 of 2
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 3293)
AUTHORS Goetz,M., Godt,D.E., Guivarc'h,A., Kahmann,U., Chriqui,D. and
Roitsch,T.
TITLE Induction of male sterility in plants by metabolic engineering of
the carbohydrate supply
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (2001) In press
REFERENCE 2 (bases 1 to 3293)
AUTHORS Goetz,M., Godt,D.E. and Roitsch,T.
TITLE Direct Submision
JOURNAL Submitted (04-MAY-2001) Plant Physiology and Cell Biology,
University of Regensburg, Universitaetsstr. 31, Regensburg 93053,
Germany
FEATURES
source
1..3293
/organism="Nicotiana tabacum"
/mol_type="genomic DNA"
/db_xref="taxon:4097"
1..3290
/gene="Nin8"
<3291..>3293
/gene="Nin8"
/product="extracellular invertase Nin8"
3291..>3293
/gene="Nin8"
/codon_start=1
/product="extracellular invertase Nin8"
/protein_id="AAK57503.1"
/db_xref="GI:14211754"
/translation="M"
ORIGIN
Query Match      8.0%; Score 72.4; DB 8; Length 3293;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCCGCCGAATTAACCGACCAATTTTGGTAGTGTATTAATTTTAA 366
    |||||
Db 2819 TCAGAAATTCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760
    |||||

QY 367 ---TTTATTTTAAATGAAAACCTACCAAGTTAGTCGGTTCTTGAACATTAATTCG 423
    |||||
Db 2759 ATTTTATTTTATTTTACGAAACCGACCACTTTGGTCGGTTTTTCTTGGCGCAAAATGC 2700
    |||||

QY 424 CGGGACTCAAAAATAGTTTCCCGCATTTTTCGCCCAAGAAAACCGACCAAGTTGGTCG 483
    |||||
Db 2699 GCGAACTATTTTTCAGTCCCGCGAAATTTATGTTTCAAGAACCGACTAATTGGTTA 2640
    |||||

QY 484 GTTTCGTAAAAAATAAATAATTTAAAAATATATTTTAAAAAACCGACCACTTAGTC 543
    |||||
Db 2639 GTTTTCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2584
    |||||

QY 544 GGTTTTGGTCGATTTTGTGA 565
    |||||
Db 2583 GGTTAATTCGGCGGATCAATTA 2562
    |||||

RESULT 8
AX068012/c AX068012 3294 bp DNA linear PAT 19-JAN-2001
DEFINITION Sequence 1 from Patent WO0077187.

```

```

ACCESSION AX068012
VERSION AX068012.1 GI:12329820
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 007187-A 1 21-DEC-2000;
Roitsch, Thomas, Dr. (DE)
FEATURES
source
1..3294
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Query Match 8.0%; Score 72.4; DB 6; Length 3294;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCCCGCAATTAACCGACCAATTTTGGTAGGTTTCTTAAATATTTTAA- 366
DB 2819 TCAGAAATTCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760

QY 367 ---TTTATTAAATGAAAAAATAACCAAGTTAGTCGGTTTCTTGAACAATAAATTCG 423
DB 2759 ATTATTTTATTTTACGAAACCGACCAACTTTGGTCGGTTTTCTTGGCGCAAAATGC 2700

QY 424 CGGACTCAAAAATAGTTTCCGGCATTTTGGCCCAAGAAAACCGACCAAGTTGGTCG 483
DB 2699 GGGAAACTATTTTGGAGTCCCGCAATTTATGTTTCAAGAAACCGACTAATTTGGTTA 2640

QY 484 GTTTCGTAAAAAATAAATAATTTAAAAATATATTTAAAAAACCAGCAACTTTAGTC 543
DB 2639 GTTTTCAATTAATAATAA---TAAAAATTAATTAATAAACCCTACCAAAATGGTC 2584

QY 544 GGTATTTGGTCGATTTTGA 565
DB 2583 GGTAAATTCGGCGGATCATTTA 2562

RESULT 9
LOCUS AX068019/c
DEFINITION Sequence 8 from Patent WO007187.
ACCESSION AX068019
VERSION AX068019.1 GI:12329827
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 007187-A 8 21-DEC-2000;
Roitsch, Thomas, Dr. (DE)
FEATURES
source
1..4135
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Fusion aus Promotor und codierender Sequenz in antisense-Orientierung"
ORIGIN
Query Match 8.0%; Score 72.4; DB 6; Length 4135;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCCCGCAATTAACCGACCAATTTTGGTAGGTTTCTTAAATATTTTAA- 366
DB 3837 TCAGAAATTCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 3778

QY 367 ---TTTATTAAATGAAAAAATAACCAAGTTAGTCGGTTTCTTGAACAATAAATTCG 423
DB 3777 ATTATTTTATTTTACGAAACCGCAACTTTGGTCGGTTTTCTTGGCGCAAAATGC 3718

QY 424 CGGACTCAAAAATAGTTTCCGGCATTTTGGCCCAAGAAAACCGACCAAGTTGGTCG 483
DB 3717 GGGAAACTATTTTGGAGTCCCGCAATTTATGTTTCAAGAAACCGACTAATTTGGTTA 3658

QY 484 GTTTCGTAAAAAATAAATAATTTAAAAATATATTTAAAAAACCAGCAACTTTAGTC 543
DB 3657 GTTTTCAATTAATAATAA---TAAAAATTAATTAATAAACCCTACCAAAATGGTC 3602

QY 544 GGTATTTGGTCGATTTTGA 565
DB 3601 GGTAAATTCGGCGGATCATTTA 3580

RESULT 11
LOCUS AC090485/c
DEFINITION Sequence 2 from Patent WO007187.
ACCESSION AX068013
VERSION AX068013.1 GI:12329821
KEYWORDS Nicotiana tabacum (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 007187-A 2 21-DEC-2000;
Roitsch, Thomas, Dr. (DE)
FEATURES
Location/Qualifiers
1..4312
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Query Match 8.0%; Score 72.4; DB 6; Length 4312;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCCCGCAATTAACCGACCAATTTTGGTAGGTTTCTTAAATATTTTAA- 366
DB 3837 TCAGAAATTCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 3778

QY 367 ---TTTATTAAATGAAAAAATAACCAAGTTAGTCGGTTTCTTGAACAATAAATTCG 423
DB 3777 ATTATTTTATTTTACGAAACCGCAACTTTGGTCGGTTTTCTTGGCGCAAAATGC 3718

QY 424 CGGACTCAAAAATAGTTTCCGGCATTTTGGCCCAAGAAAACCGACCAAGTTGGTCG 483
DB 3717 GGGAAACTATTTTGGAGTCCCGCAATTTATGTTTCAAGAAACCGACTAATTTGGTTA 3658

QY 484 GTTTCGTAAAAAATAAATAATTTAAAAATATATTTAAAAAACCAGCAACTTTAGTC 543
DB 3657 GTTTTCAATTAATAATAA---TAAAAATTAATTAATAAACCCTACCAAAATGGTC 3602

QY 544 GGTATTTGGTCGATTTTGA 565
DB 3601 GGTAAATTCGGCGGATCATTTA 3580

```

```

DEFINITION Genomic Sequence for Oryza sativa, Nipponbare strain, clone
ACCESSION OSJNBa0067N01, from chromosome 3, complete sequence.
VERSION AC090485.3 GI:14495364
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 159636)
AUTHORS de la Bastide,M., Spiegel,L., Preston,R., Kirchoff,K., King,L.,
Nascimento,L., Baker,J., Vil,M.D., Zutavern,T., Santos,L.,
Miller,B., Kuit,K., Cunnius,D.M., Rodriguez,S., Balija,V., Shah,R.,
Bahret,A., Yang,C., Bell,M., Palmer,L., O'Shaughnessy,A., Dedhia,N.
and McCombie,W.R.
TITLE Genomic Sequence for Oryza sativa, Nipponbare strain, clone
JOURNAL OSJNBa0067N01, from chromosome 3, complete sequence
REFERENCE 2 (bases 1 to 159636)
AUTHORS McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
NY 11724, USA
REFERENCE 3 (bases 1 to 159636)
AUTHORS McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2001) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
NY 11724, USA
REFERENCE 4 (bases 1 to 159636)
AUTHORS Palmer,L.E., de la Bastide,M., Spiegel,L., Preston,R., Kirchoff,K.,
King,L., Nascimento,L., Baker,J., Vil,M.D., Zutavern,T., Santos,L.,
Miller,B., Kuit,K., Cunnius,D.M., Rodriguez,S., Balija,V., Shah,R.,
Bahret,A., Yang,C., Bell,M., O'Shaughnessy,A., Dedhia,N. and
McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2001) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY
11724, USA
REMARK Genomic Sequence for Oryza sativa, Nipponbare strain, clone
OSJNBa0067N01, from chromosome 3, complete sequence
COMMENT On Jun 20, 2001 this sequence version replaced gi:14209722.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.
FEATURES
source
1. 159636
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0067N01"
/clone.lib="HindIII"
/complement(851..931)
/note="transposon02 TNP2-like transposon protein gi|4"
1532..2626
/gene="OSJNBa0067N01.1"
/note="Hypothetical protein"
1532..2626
/gene="OSJNBa0067N01.1"
/note="Similar to Pir||T46123 hypothetical protein
T2013.90 - Arabidopsis thaliana identities = 119/339
Evalue= 3e-44"
/codon_start=1
/product="Hypothetical protein"
/protein_id="AAK98722.1"
/db_xref="GI:15451601"

/db_xref="GI:15451598"
/translation="MIDAVHPKVKVAPSSSFDFREYTSAFRTESYNDVFWVLDTIT
LAHGAALVFRHGGGCGCAARLFSYRLFADHLLPEQORAAVALASGRGRLPDVR
GLLAAYAEATANAFSLCSHLKLDIEHLRRLFKTKRLKLSARVSLGVSLADVSAAALG
QPFTALAAOQRLREVQAGDLRLGLDAGRKARHRSVARLRSALVSFVFAVAV
VAVGACIGVHILAAFAAPPMWGSFGRPAAGGRRALVQLAAAKGTIILNRDM
ETISRLVARVRDEGHMVALLRLCVRHPAAGAGKGRVLQVLELRQLSKNEESFRQL
DELEBHLFCFMTTKARIMYMFMAAAA"
/complement(3166..5556)
/gene="OSJNBa0067N01.2"
/note="Putative signal peptidase"
3166..5452,5534..5556)
/gene="OSJNBa0067N01.2"
/note="Similar to emb|CAC09470.1| (AL442113) putative
single peptidase [Oryza sativa] Identities = 68/138
Evalue= 7e-33"
/codon_start=1
/product="Putative signal peptidase"
/protein_id="AAK98723.1"
/db_xref="GI:15451599"
/translation="MENGNLCELELGAVLAEBSCLQYQFSDVDVLFKCPDHRRE
LFRKLIAPGEMWQLPGTPIIKIPGHCWVEGDNAACSDRSRSGFEVVGIDKDSMG
GVRYSASAGMIGPPRIPLUGLIGRVAHVWPPSKIGRVDTKPENRISPL"
7459..10855
/gene="OSJNBa0067N01.3"
/note="Unknown protein with similarity to
methyltransferases and transcriptional regulators"
Join(7459..9132,9247..9330,9641..9745,9839..9947,
10446..10597,10718..10855)
/gene="OSJNBa0067N01.3"
/note="Similar to dbj|BAA88534.1| (AP000969) Similar to
Arabidopsis thaliana DNA chromosome 4, BAC clone T5L19
(AL049481) [Oryza sativa] Identities = 571/759 Evalue=
0.0. Similar to sp|P25583|KAR4 YEAST KARYOGAMY PROTEIN
KAR4 Identities = 98/328 Evalue= 2e-27"
/codon_start=1
/product="Unknown protein with similarity to
methyltransferases and transcriptional regulators"
/protein_id="AAK98724.1"
/db_xref="GI:15451600"
/translation="MDSRPKIPRSPPDSKDKSDNRGRDEKNDVSSRTYGETD
CKEBCDNRKKGAMSGDVGDSDRSVDRSHETVHVNDKQKAVEIKNLIHLDGVGQ
SDYQQLGDLNERNRNGTVKSRDAHDDKLGSRDRNWTGKQTPGSGVDYLRKSC
SQDSKASDSESKNAQERQDGGFGRVGRDFRGRSESTRGSTGYDTSIE
IRPNNSLDFREGSVSGRYDVGVAHRDVTYTGNDKVTNSEPQSGSASISQFPQH
CHKGRPSRGGRPNRSDORSQVTLPIPMPPFGLPLGPPGPHIGPNIPSPGPH
LLPGVFVPPFGGPLLWPGAGVDVNLSPNLPVPPVAGEHSTFGMGAGPNHIL
NQFGSIGAPTNMAGLSLPHQLGTSREMHVCKPPVGGGTNRNMSGPTKAGSREON
DYSQNFVDTGRPNQPIRELDLTSVADYFKRELIIQKDEIVAKSASPFWYKCDLR
QVLSPEFFGTFKFDVLVDPPWRYVHRAPGITDHIYXWNGEEMNLIKIEAIDTSPF
VLMVGDGGLQGRCLKQFRCEDEVCKVTKKSAFSLRSHDSTIILOHKEHC
LMGIKGTVRSTDGHVHANIDTIIIADEPTDGTSTKPPEDMYRIIEHFALGKRLEL
FGEDHNIRPGWLTILGKLSYNFNKAYVKNFADKDGKVGQGRNPPPEAPHLVVTP
RIEGLRPSKPSHK"
14781..14815
/note="misc_rep03 similar to Rat tandem repeat mRNA"
/complement(16651..16787)
/note="Galign_Ol2 MITE element from gb:U72728 Oryza 1"
/complement(17170..17328)
/note="Wanderer_Os5 MITE element from gb:X13679 Oryza"
19417..20916
/gene="OSJNBa0067N01.4"
/note="Putative retroelement"
Join(19417..20491,20573..20916)
/gene="OSJNBa0067N01.4"
/note="Similar to gb|AAK38301.1|AC079029.2 (AC079029)
putative reverse transcriptase [Oryza sativa] Identities =
170/488 Evalue= 3e-59"
/codon_start=1
/product="Putative retroelement"
/protein_id="AAK98725.1"
/db_xref="GI:15451601"

repeat_region
repeat_region
repeat_region
gene
CDS

```



```

DEFINITION L.luteus tRNA-Gly genes and tRNA-Gly pseudogenes.
ACCESSION Z49255
VERSION Z49255.1 GI:2073451
KEYWORDS pseudogene; transfer RNA-Gly; tRNA-Gly gene.
SOURCE Lupinus luteus (yellow lupine)
ORGANISM Lupinus luteus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae;
Lupinus.
1
REFERENCE 1
AUTHORS Nuc,P., Nuc,K., Szweykowska-Kulinska,Z. and Pawelkiewicz,J.
TITLE Nucleotide sequence of nuclear tRNA(Gly) genes and tRNA(Gly) pseudogenes from yellow lupin (Lupinus luteus): expression of the tRNA(Gly) genes in vitro and in vivo
JOURNAL Acta Biochim. Pol. 44 (2), 259-274 (1997)
MEDLINE 98025190
PUBMED 9360715
REFERENCE 2 (bases 1 to 7869)
AUTHORS Nuc,P.W.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1995) Nuc P. W., University of Agriculture, Biochemistry & Biotechnology, ul.Wolynska 35, Poznan, Poland, 60 637
COMMENT On May 9, 1997 this sequence version replaced gi:1924956.
FEATURES             Location/Qualifiers
     source            1..7869
                     /organism="Lupinus luteus"
                     /mol_type="genomic DNA"
                     /cultivar="ventus, line no. 098067, Poland, Wiatrowo, W.Swiecicki"
                     /db_xref="taxon:3873"
                     /clone="lambda G4"
                     /tissue_type="epicotyl"
                     /clone_lib="EMBL3 genomic library"
                     /dev_stage="seedling"
     repeat_region     323..6668
                     /rpt_type=TANDEM
                     /size=332
     gene              332..402
                     /gene="tRNA-Gly"
     tRNA              332..402
                     /gene="tRNA-Gly"
                     /product="tRNA-Gly"
                     /note="codon recognized: GGC"
                     /anticodon=(pos:364..366,aa:Gly)
                     /evidence=experimental
     gene              1075..1145
                     /gene="tRNA-Gly"
     tRNA              1075..1145
                     /gene="tRNA-Gly"
                     /product="tRNA-Gly"
                     /pseudo
     gene              2897..2967
                     /gene="tRNA-Gly"
     tRNA              2897..2967
                     /gene="tRNA-Gly"
                     /product="tRNA-Gly"
                     /evidence=experimental
     gene              3639..3709
                     /gene="tRNA-Gly"
     tRNA              3639..3709
                     /gene="tRNA-Gly"
                     /product="tRNA-Gly"
                     /pseudo
     gene              5461..5531
                     /evidence=experimental
     tRNA              5461..5531
                     /gene="tRNA-Gly"
                     /product="tRNA-Gly"
                     /evidence=experimental
     gene              6201..6271
                     /gene="tRNA-Gly"

```

```
tRNA      6201.. 6271
/gene="tRNA-Gly"
/product="tRNA-Gly"
/pseudo

ORIGIN
Query Match      7.6%; Score 68.8; DB 8; Length 7869;
Best Local Similarity 59.0%; Pred. No. 0.0057;
Matches 118; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 36 TTAATTCAAAATAAATTTGTAATAAAACCGACCAAGTTGATCGGTATTTAATTA 95
DB 5604 TTCATCTTAGAGATGTTAGTGAATGACCAACAAACACATATATGCTAATATCATCATAG 5545
QY 96 TGTATATAAAGATTTCATCTCTCGGAATCGAACCGGGTCTGACTATGCGAAGATACT 155
DB 5544 CTCGAAGTATATATGACACCGCGGGAATCGAACCGGGTCTGACTATGCGAAGATACT 5485
QY 156 ATTCACCACTAGACCAATGTTGTTTATTTGTTTAAAGACTGTCCTTTTATTGATTATAC 215
DB 5484 ATTCACCACTAGACCACTGTTGTTTATTTGTTTAAAGACTGTCCTTTTATTGATTATAC 5425
QY 216 TCTTTATATTATTTTGTCA 235
DB 5424 TAAATGCTTATATTCTTACA 5405

RESULT 15
NAU40606/c
LOCUS      NAU40606      628 bp      DNA      linear      PLN 23-AUG-2001
DEFINITION Nicotiana alata isolate S2S2 S-like RNase gene, intron sequence.
ACCESSION  U40606
VERSION     U40606.1  GI:1272238
KEYWORDS
SOURCE      Nicotiana alata (Persian tobacco)
ORGANISM    Nicotiana alata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 628)
            Dodds,P.N., Clarke,A.E. and Newbiggin,E.
            Molecular characterisation of an S-like RNase of Nicotiana alata
            that is induced by phosphate starvation
            Plant Mol. Biol. 31 (2), 227-238 (1996)
JOURNAL
MEDLINE     96343927
PUBMED      8756589
REFERENCE   2  (bases 1 to 628)
            Dodds,P.N.
            Direct Submission
            Submitted (14-NOV-1995) Peter N Dodds, Botany, University of
            Melbourne, Parkville, Victoria, 3052, Australia
FEATURES
            source
            1..628
                /organism="Nicotiana alata"
                /mol_type="genomic DNA"
                /isolate="S2S2"
                /db_xref="taxon:4087"
            1..628
                /gene="S-like RNase gene"
            1..628
                /gene="S-like RNase gene"
            275..481
                /gene="S-like RNase gene"
                /note="insertion sequence found in the intron region
                represents an allelic variant"

gene
intron
misc_feature

ORIGIN
Query Match      7.5%; Score 68; DB 8; Length 628;
Best Local Similarity 70.1%; Pred. No. 0.011;
Matches 103; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 600 TTTCGCGAATTTCTAGTAGTACCGAACCCTGTAAGCTTCGGAGAAAATTTGTATATCT 659
```

```
Db      419 TATGTCTAGTATCAAGGGATTACCGCACACCGTAAACTTCGCATAAACCTTGTATATGT 360
QY      660 ATATGTGTATATCCTTTAAATGATTAAATTTAAAGAACGNNGCACCCCTGAATACTAGAAGC 719
Db      359 TTAACCTACATACCTTAAAAAATGATT-ATATAAATAAATTTGGCACCCCTAAACGTTAAAAANC 301
QY      720 CTTTAGGGGCACCTAGATGAGCAGAATA 746
Db      300 CTTTAGGGGTACTGAGTGAGCAAAATA 274

Search completed: November 11, 2004, 10:27:04
Job time : 3881.86 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 06:24:54 ; Search time 444.342 Seconds
(without alignments)
10644.348 Million cell updates/sec

Title: US-10-009-966c-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tgggtcggaattaccgaccca.....atgacttcattcgtattctc 901

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	899	99.8	3294	4 AAF25735	Aaf25735 Tobacco i
2	899	99.8	4135	4 AAF25742	Aaf25742 Tobacco p
3	899	99.8	4312	4 AAF25736	Aaf25736 Tobacco i
C 4	72.4	8.0	3294	4 AAF25735	Aaf25735 Tobacco i
C 5	72.4	8.0	4135	4 AAF25742	Aaf25742 Tobacco p
C 6	72.4	8.0	4312	4 AAF25736	Aaf25736 Tobacco i
C 7	64.8	7.2	1140	10 ADB87452	Adb87452 Transgene
C 8	64.8	7.2	1140	10 ADB87474	Adb87474 Transgene
C 9	64.8	7.2	1140	12 ADJ35261	Adj35261 Thale cre
C 10	64.8	7.2	1140	12 ADJ35283	Adj35283 Thale cre
C 11	64.8	7.2	1140	12 ADL97593	Adl97593 Stabilizin
C 12	64.8	7.2	1140	12 ADL96879	Adl96879 Stabilizin
C 13	64	7.1	5945	6 ABL32085	Ab132085 Human imm
C 14	62.8	7.0	15518	6 ABL34172	Ab134172 Human imm
C 15	62.8	7.0	15518	6 ABL34624	Ab134624 Human met
C 16	62.8	7.0	15518	6 ABL70607	Ab170607 Chemical
C 17	61.8	6.9	91608	10 AAL54538	Aal54538 Arabidops
C 18	57	6.3	6816	12 ADQ24856	Adq24856 Human sof
C 19	55.4	6.1	6239	6 ABL33776	Ab133776 Human imm
C 20	55.4	6.1	6239	6 ABK28297	Abk28297 DNA trans
C 21	55.2	6.1	5407	6 ABL34091	Ab134091 Human imm

22	54.8	6.1	8056	8 ABZ10246	Abz10246 Haematopo
C 23	54.6	6.1	368	4 AAS60046	Aas60046 Human can
24	53.4	5.9	8056	8 ABZ10100	Abz10100 Haematopo
25	53	5.9	6092	6 AAS61080	Aas61080 Human gen
C 26	52.8	5.9	8979	6 ABL32784	Ab132784 Human imm
C 27	52.8	5.9	8979	6 ABK31270	Abk31270 Signal tr
C 28	52.8	5.9	8979	6 ABL70231	Ab170231 Chemical
C 29	52.8	5.9	8979	6 ABL61178	Abx61178 Human gen
C 30	52.4	5.8	419	8 ABX46069	Abx46069 Bovine ES
C 31	52	5.7	2000	8 ABL55643	Ab155643 AnEPV gen
C 32	51.4	5.7	2000	8 ADA71938	Ada71938 Rice gene
C 33	51.2	5.7	358	4 AAI83451	Aai83451 Human pol
C 34	51	5.7	16914	6 ABL70316	Ab170316 Chemical
C 35	51	5.7	16914	6 AAS61254	Aas61254 Human gen
C 36	50.4	5.6	5756	6 ABL32586	Ab132586 Human imm
C 37	50.4	5.6	7234	6 ABK31228	Abk31228 Signal tr
C 38	50.4	5.6	15732	4 AAS45389	Aas45389 Chemical
C 39	50.4	5.6	15732	6 ABK28234	Abk28234 DNA trans
C 40	50.2	5.6	500	5 ADI73089	Adi73089 Human ova
C 41	50.2	5.6	500	5 ADL38222	Adl38222 Human ova
C 42	50	5.5	5997	6 ABL33625	Ab133625 Human imm
C 43	49.6	5.5	417	5 ABV09124	Abv09124 Human pro
C 44	49.6	5.5	1501	8 ABZ10042	Abz10042 Haematopo
C 45	49.6	5.5	1501	10 ADE84086	Ade84086 Human lym

ALIGNMENTS

RESULT 1
AAF25735
ID AAF25735 standard; DNA; 3294 BP.

AC AAF25735;

DT 06-APR-2001 (first entry)

DE Tobacco invertase promoter SEQ ID NO 1.

KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.

OS Nicotiana tabacum.

PN WO200007187-A2.

PD 21-DEC-2000.

PF 13-JUN-2000; 2000WO-DE001944.

PR 12-JUN-1999; 99DE-02009998.

PR 04-APR-2000; 2000DE-02005992.

PR 26-APR-2000; 2000DE-02007494.

PA (ROIT/) ROITSCH T.

PI Roitsch T;

PS WPI; 2001-080685/09.

XX New tapetum- and pollen-specific promoter from tobacco, useful for
XX preparing male sterile plants, particularly those with seedless fruits.

XX Claim 3; Page 64-65; 74pp; German.

XX This invention describes a novel nucleic acid (I) comprising a promoter
XX that is specific for tapetum and pollen. The invention also describes (a)
XX expression system containing at least one (i); (b) nucleic acid construct
XX (ii) containing (i) plus at least part of an expressible nucleic acid
XX (iii); (c) vector containing (i), the system of (a), or (ii); (d) cells,
XX particularly plant cells, containing (i), the system of (a), (ii) or the
XX vector of (c); (e) plants containing cells of (d); (f) seeds from the

parent
PCT

DB 4113 TAAACATAAGACTCGTCTTTATAGCACCTTTTAAATAAGACTATGACTTCATCTGATTTCT 41/2

XX
XX
XXXXXXXXXXXX

Sequence 3294 BP; 983 A; 597 C; 644 G; 1064 T; 0 U; 6 Other;

```

Query Match      8.0%; Score 72.4; DB 4; Length 3294;
Best Local Similarity 60.3%; Pred. No. 3.2e-05;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCGCGCAATTAAACGACCAATTTTGGTAGGTTTTTTTAAATATTTTAA- 366
DB 2819 TCAAAAATCGACCAAAAACGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760

QY 367 ---TTTATTTTAAATGAAACTAACCAAGTTAGTCGGTTCTTGAACATAAATTCG 423
DB 2759 ATTTTATTTTATTTTACGAAACGACCAACTTTGGTCGGTTTTCTTTGGCGCAAAATGC 2700

QY 424 CGGAGCTCAAAAATAGTTTCCCGCATTTTTCGCGCAAAAGAAACCGACCAAAAGTTGGTCG 483
DB 2699 GGGAACTATTTTGGAGTCGCGGAATTTATGTTTCAAGAAACGACTAACTTTGGTTA 2640

QY 484 GTTTCGTAAAAAATAAATTTTAAATAATATATTTTAAATAAATTTTAAATAAATTTAGTC 543
DB 2639 GTTTTCAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATGGTC 2584

QY 544 GGTATTTTGGTCGATTTTTCGA 565
DB 2583 GGTAAATTCGGCGGATCATTTA 2562

RESULT 5
AAF25742/c
ID AAF25742 standard; DNA; 4135 BP.
XX
AC AAF25742;
DT
DT 06-APR-2001 (first entry)
XX
DE Tobacco promoter/invertase NIN88 DNA fragment SEQ ID NO 8.
XX
KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.
XX
OS Nicotiana tabacum.
XX
PN WO200077187-A2.
XX
PD 21-DEC-2000.
XX
PF 13-JUN-2000; 2000WO-DE001944.
XX
PR 12-JUN-1999; 99DE-02009998.
XX
PR 04-APR-2000; 2000DE-02005992.
XX
PR 26-APR-2000; 2000DE-02007494.
XX
PA (ROIT/) ROITSCH T.
XX
PI Roitsch T;
XX
PI WPI; 2001-080685/09.
XX
DR
XX
PT New tapetum- and pollen-specific promoter from tobacco, useful for
PT preparing male sterile plants, particularly those with seedless fruits.
XX
PS Example 4.3; Page 68-70; 74pp; German.
XX
CC This invention describes a novel nucleic acid (I) comprising a promoter
CC that is specific for tapetum and pollen. The invention also describes (a)
CC expression system containing at least one (i); (b) nucleic acid construct
CC (iii) containing (i) plus at least part of an expressible nucleic acid
CC (iii); (c) vector containing (i), the system of (a), or (ii); (d) cells,
CC particularly plant cells, containing (i), the system of (a), (ii) or the
CC vector of (c); (e) plants containing cells of (d); (f) seeds from the
CC plants of (e); (g) hybrid seed produced by crossing a male sterile plant
CC of (e); (h) production of male sterile plants by introducing (ii) into a
CC cell and regeneration to a plant; (i) restorer plants containing in one

```

```

(or preferably many) cells a construct (IIa) of (i) and a sequence that
encodes an invertase different from the plants endogenous invertase; (j)
plant that contains in one (or preferably many) cells both (ii) and (IIa)
; (k) seeds from plants of (i) and (j); (l) fruits, particularly
seedless, produced by plants of (e), (i) and (j); and (m) method for
cloning a promoter that is functionally homologous with (i). Constructs
containing (i) and an invertase-encoding sequence are used to produce
male-sterile plants (by co-suppression or antisense techniques) for
preparation of hybrids, while constructs that contain (i) and a
heterologous invertase-encoding sequence are used to prepare restorer
plants (which allow propagation of the male-sterile plants). Seeds from
these plants are used for in vitro embryogenesis of haploid or (double)
diploid plants, and the new plants particularly produce seedless fruits.
(i) can also be used to prepare transgenic plants that show increased or
reduced production of endogenous materials, e.g. of plant hormones or
proteins involved in provision of energy to developing tissue. (i)
provide high level expression in a tissue- and time-specific manner, and
do not require exogenous stimuli (e.g. temperature or chemicals) for
regulation
XX
SQ Sequence 4135 BP; 1212 A; 774 C; 825 G; 1318 T; 0 U; 6 Other;

Query Match      8.0%; Score 72.4; DB 4; Length 4135;
Best Local Similarity 60.3%; Pred. No. 3.3e-05;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCGCGCAATTAAACGACCAATTTTGGTAGGTTTTTTTAAATATTTTAA- 366
DB 2819 TCAAAAATCGACCAAAAACGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760

QY 367 ---TTTATTTTAAATGAAACTAACCAAGTTAGTCGGTTCTTGAACATAAATTCG 423
DB 2759 ATTTTATTTTATTTTACGAAACGACCAACTTTGGTCGGTTTTCTTTGGCGCAAAATGC 2700

QY 424 CGGAGCTCAAAAATAGTTTCCCGCATTTTTCGCGCAAAAGAAACCGACCAAAAGTTGGTCG 483
DB 2699 GGGAACTATTTTGGAGTCGCGGAATTTATGTTTCAAGAAACGACTAACTTTGGTTA 2640

QY 484 GTTTCGTAAAAAATAAATTTTAAATAATATATTTTAAATAAATTTTAAATAAATTTAGTC 543
DB 2639 GTTTTCAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATGGTC 2584

QY 544 GGTATTTTGGTCGATTTTTCGA 565
DB 2583 GGTAAATTCGGCGGATCATTTA 2562

RESULT 6
AAF25736/c
ID AAF25736 standard; DNA; 4312 BP.
XX
AC AAF25736;
DT
DT 06-APR-2001 (first entry)
XX
DE Tobacco invertase promoter SEQ ID NO 2.
XX
KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.
XX
OS Nicotiana tabacum.
XX
PN WO200077187-A2.
XX
PD 21-DEC-2000.
XX
PF 13-JUN-2000; 2000WO-DE001944.
XX
PR 12-JUN-1999; 99DE-02009998.
XX
PR 04-APR-2000; 2000DE-02005992.
XX
PR 26-APR-2000; 2000DE-02007494.
XX

```


OS Unidentified.
XX WO2003004704-A2.
XX 16-JAN-2003.
XX 14-JUN-2002; 2002WO-NL000390.
XX 04-JUL-2001; 2001EP-0020581.
XX 05-JUL-2001; 2001US-0303199P.
XX (CHRO-) CHROMAGENICS BV.
XX Otte AP, Kruckeberg AL;
XX WPI; 2003-229412/22.
XX Selecting a DNA sequence with a gene transcription modulating quality by
XX providing a transcription system with a variety fragment-comprising
XX vectors and performing a selection step in the transcription system.
XX Claim 43; Fig 26; 216pp; English.
XX The invention relates to DNA sequences with gene transcription regulatory
XX qualities and methods for the detection and use of the regulatory DNA
XX sequences. The invention further comprises providing a transcription
XX system with a variety fragment-comprising vectors; and performing a
XX selection step in the transcription system in order to identify a
XX fragment comprising the DNA sequence with the gene transcription
XX modulating activity. This polynucleotide represents a STAR element used
XX in the method of the invention. The STAR element is shown to improve
XX transgene expression.
XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
SQ
Query Match 7.2%; Score 64.8; DB 10; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAAGATTTCACCTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 157
DB 366 TCATAGCTCAATGCACGCGGGATCGAACCGGGTCTGTACTATGGCAAGATACTAT 307
QY 158 TCTACCACTAGACCATGGTTCATTTTGTAAAGACTGCTTTTATTTGATTATATCTC 217
DB 306 TCTACCACTAGACCATGGTTCATTTTGTAAAGACTGCTTTTATTTGATTATATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATACCGAACCAAGTTGGTCGATTTTATTTAAAGTA 277
DB 246 ATTATAAAGTATCATGTATTAACACACCAATCTTTTGTAGTTTGTACCACTACGAAGTA 187
RESULT 9
ADJ35261/c
ID ADJ35261 standard; DNA; 1140 BP.
XX AC ADJ35261;
XX 22-APR-2004 (first entry)
XX Thale cress stabilising anti-repression, STAR, element #13.
XX STAR affiliated proteinaceous molecule; post translational modification;
XX thale cress; stabilising anti-repression; STAR; STAR element; ds.
XX Arabidopsis thaliana.
XX WO2003106674-A2.
XX 24-DEC-2003.
XX 30-MAY-2003; 2003WO-NL000410.
XX

PR 14-JUN-2002; 2002EP-00077344.
XX (CHRO-) CHROMAGENICS BV.
XX Otte AP, Kruckeberg AL, Satijn DPE;
XX WPI; 2004-082195/08.
XX Producing proteinaceous molecules in cells by selecting a cell, providing
XX a nucleic acid encoding a proteinaceous molecule with an Stabilizing Anti
XX -Repression sequence and expressing proteinaceous molecule.
XX Example 13; SEQ ID NO 97; 177pp; English.
XX The invention relates to a method of producing a proteinaceous molecule
XX (I) in a cell comprising selecting a cell for its suitability for
XX producing (I), providing a nucleic acid encoding (I) with a nucleic acid
XX comprising a Stabilising Anti-Repression (STAR) sequence, expressing the
XX resulting nucleic acid in the cell and collecting (I). The method is
XX useful for producing (I). A cell line (II) provided with a nucleic acid
XX comprising a STAR sequence is useful for producing (I). (II) Enables
XX production of affiliated proteinaceous molecule, as cell carries out
XX proper post-translational modifications of produced proteins. The present
XX sequence represents a thale cress stabilising anti-repression, STAR,
XX element.
XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
SQ
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAAGATTTCACCTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 157
DB 366 TCATAGCTCAATGCACGCGGGATCGAACCGGGTCTGTACTATGGCAAGATACTAT 307
QY 158 TCTACCACTAGACCATGGTTCATTTTGTAAAGACTGCTTTTATTTGATTATATCTC 217
DB 306 TCTACCACTAGACCATGGTTCATTTTGTAAAGACTGCTTTTATTTGATTATATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATACCGAACCAAGTTGGTCGATTTTATTTAAAGTA 277
DB 246 ATTATAAAGTATCATGTATTAACACACCAATCTTTTGTAGTTTGTACCACTACGAAGTA 187
RESULT 10
ADJ35283/c
ID ADJ35283 standard; DNA; 1140 BP.
XX AC ADJ35283;
XX 22-APR-2004 (first entry)
XX Thale cress stabilising anti-repression, STAR, element #35.
XX STAR affiliated proteinaceous molecule; post translational modification;
XX thale cress; stabilising anti-repression; STAR; STAR element; ds.
XX Arabidopsis thaliana.
XX WO2003106674-A2.
XX 24-DEC-2003.
XX 30-MAY-2003; 2003WO-NL000410.
XX 14-JUN-2002; 2002EP-00077344.
XX (CHRO-) CHROMAGENICS BV.
XX Otte AP, Kruckeberg AL, Satijn DPE;
XX WPI; 2004-082195/08.
XX

XX Producing proteinaceous molecules in cells by selecting a cell, providing
PT a nucleic acid encoding a proteinaceous molecule with an Stabilizing Anti-
PT -Repressor sequence and expressing proteinaceous molecule.
XX Example 13; SEQ ID NO 119; 177pp; English.
XX The invention relates to a method of producing a proteinaceous molecule
CC (I) in a cell comprising selecting a cell for its suitability for
CC producing (I), providing a nucleic acid encoding (I) with a nucleic acid
CC comprising a Stabilizing Anti-Repressor (STAR) sequence, expressing the
CC resulting nucleic acid in the cell and collecting (I). The method is
CC useful for producing (I). A cell line (II) provided with a nucleic acid
CC comprising a STAR sequence is useful for producing (I). (II) Enables
CC production of affiliated proteinaceous molecule, as cell carries out
CC proper post-translational modifications of produced proteins. The present
CC sequence represents a thale cress stabilising anti-repression, STAR,
CC element.
XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
SQ
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 157
Db 366 TCATAGCTCAATGCACCGCCGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 307
QY 158 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATCT 217
Db 306 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATACCGACCAAGTTGTCGATTTTATTAAGAAGTA 277
Db 246 ATTATAAAGTATCATGTATAAACAACAACAATCTTTTGTAGTTTGTACCACTACGAAGTA 187
RESULT 11
ADL97593/c
ID ADL97593 standard; DNA; 1140 BP.
XX AC ADL97593;
XX 20-MAY-2004 (first entry)
XX Stabilizing Anti-Repressor DNA sequence, A35.
XX protein expression unit; Stabilizing Anti-Repressor; STAR; heterologous;
XX multimeric; transgene expression; ds.
XX Arabidopsis sp.
XX WO2003106684-A2.
XX 24-DEC-2003.
XX 13-JUN-2003; 2003WO-NL000432.
XX 14-JUN-2002; 2002EP-00077350.
XX (CHRO-) CHROMAGENICS BV.
XX Otte AP, Kruckeberg AL, Sewalt RGAB;
XX WPI; 2004-082197/08.
XX The invention relates to a novel method for obtaining a cell which
CC expresses two or more proteins. The method comprises providing the cell
CC with two or more protein expression units encoding the two or more
CC proteins, having at least two of the protein expression units with at
CC least one Stabilizing Anti-Repressor (STAR) sequence. The method provides
CC increased predictability in the production of recombinant cell lines that
CC effectively produce heterologous multimeric proteins of interest,
CC increase the yield of heterologous multimeric proteins, stable expression
CC of heterologous multimeric proteins and favorable transgene expression.
CC This polynucleotide sequence represents the DNA of a Stabilizing Anti-
CC Receptor (STAR) sequence of the invention.
XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
SQ
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 157
Db 366 TCATAGCTCAATGCACCGCCGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 307
QY 158 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATCT 217
Db 306 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATACCGACCAAGTTGTCGATTTTATTAAGAAGTA 277
Db 246 ATTATAAAGTATCATGTATAAACAACAACAATCTTTTGTAGTTTGTACCACTACGAAGTA 187
RESULT 12
ADL96879/c
ID ADL96879 standard; DNA; 1140 BP.
XX AC ADL96879;
XX 20-MAY-2004 (first entry)
XX Stabilizing Anti-Repressor DNA sequence, A13.
XX protein expression unit; Stabilizing Anti-Repressor; STAR; heterologous;
XX multimeric; transgene expression; ds.
XX Arabidopsis sp.
XX WO2003106684-A2.
XX 24-DEC-2003.
XX 13-JUN-2003; 2003WO-NL000432.
XX 14-JUN-2002; 2002EP-00077350.
XX (CHRO-) CHROMAGENICS BV.
XX Otte AP, Kruckeberg AL, Sewalt RGAB;
XX WPI; 2004-082197/08.
XX The invention relates to a novel method for obtaining a cell which
CC expresses two or more proteins. The method comprises providing the cell
CC with two or more protein expression units encoding the two or more
CC proteins, having at least two of the protein expression units with at
CC least one Stabilizing Anti-Receptor (STAR) sequence. The method provides
CC increased predictability in the production of recombinant cell lines that
CC effectively produce heterologous multimeric proteins of interest,
CC increase the yield of heterologous multimeric proteins, stable expression
CC of heterologous multimeric proteins and favorable transgene expression.
CC This polynucleotide sequence represents the DNA of a Stabilizing Anti-
CC Receptor (STAR) sequence of the invention.
XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
SQ
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 157
Db 366 TCATAGCTCAATGCACCGCCGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 307
QY 158 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATCT 217
Db 306 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATACCGACCAAGTTGTCGATTTTATTAAGAAGTA 277
Db 246 ATTATAAAGTATCATGTATAAACAACAACAATCTTTTGTAGTTTGTACCACTACGAAGTA 187

CC The invention relates to a novel method for obtaining a cell which
CC expresses two or more proteins. The method comprises providing the cell
CC with two or more protein expression units encoding the two or more
CC proteins, having at least two of the protein expression units with at
CC least one Stabilizing Anti-Repressor (STAR) sequence. The method provides
CC increased predictability in the production of recombinant cell lines that
CC effectively produce heterologous multimeric proteins of interest,
CC increase the yield of heterologous multimeric proteins, stable expression
CC of heterologous multimeric proteins and favorable transgene expression.
CC This polynucleotide sequence represents the DNA of a Stabilizing Anti-
CC Receptor (STAR) sequence of the invention.
XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
SQ
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 157
Db 366 TCATAGCTCAATGCACCGCCGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 307
QY 158 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATCT 217
Db 306 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATACCGACCAAGTTGTCGATTTTATTAAGAAGTA 277
Db 246 ATTATAAAGTATCATGTATAAACAACAACAATCTTTTGTAGTTTGTACCACTACGAAGTA 187
RESULT 12
ADL96879/c
ID ADL96879 standard; DNA; 1140 BP.
XX AC ADL96879;
XX 20-MAY-2004 (first entry)
XX Stabilizing Anti-Repressor DNA sequence, A13.
XX protein expression unit; Stabilizing Anti-Repressor; STAR; heterologous;
XX multimeric; transgene expression; ds.
XX Arabidopsis sp.
XX WO2003106684-A2.
XX 24-DEC-2003.
XX 13-JUN-2003; 2003WO-NL000432.
XX 14-JUN-2002; 2002EP-00077350.
XX (CHRO-) CHROMAGENICS BV.
XX Otte AP, Kruckeberg AL, Sewalt RGAB;
XX WPI; 2004-082197/08.
XX Obtaining (M1) a cell which expresses two or more proteins, by providing
PT cell with protein expression units encoding proteins, comprising protein
PT expression units with Stabilizing Anti-Repressor sequence.
XX Example 14; SEQ ID NO 97; 213pp; English.
XX The invention relates to a novel method for obtaining a cell which
CC expresses two or more proteins. The method comprises providing the cell
CC with two or more protein expression units encoding the two or more
CC proteins, having at least two of the protein expression units with at
CC least one Stabilizing Anti-Receptor (STAR) sequence. The method provides
CC increased predictability in the production of recombinant cell lines that
CC effectively produce heterologous multimeric proteins of interest,
CC increase the yield of heterologous multimeric proteins, stable expression
CC of heterologous multimeric proteins and favorable transgene expression.
CC This polynucleotide sequence represents the DNA of a Stabilizing Anti-
CC Receptor (STAR) sequence of the invention.
XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
SQ
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 157
Db 366 TCATAGCTCAATGCACCGCCGGGAATCGAACCGGGTCTGTACTATGGCAAGATCTAT 307
QY 158 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATCT 217
Db 306 TCTACCACTAGACCACTGGTTCATTTTGTAAAGACTGCTTTTATTGTTTATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATACCGACCAAGTTGTCGATTTTATTAAGAAGTA 277
Db 246 ATTATAAAGTATCATGTATAAACAACAACAATCTTTTGTAGTTTGTACCACTACGAAGTA 187

CC	increase the yield of heterologous multimeric proteins, stable expression	CC	diseases. The present sequence is a gene of the invention
CC	of heterologous multimeric proteins and favorable transgene expression.	XX	
CC	This polynucleotide sequence represents the DNA of a Stabilizing Anti-	SQ	Sequence 5945 BP; 2043 A; 23 C; 1010 G; 2869 T; 0 U; 0 Other;
CC	Repressor (STAR) sequence of the invention.		
XX		Query Match	7.1%; Score 64; DB 6; Length 5945;
SQ	Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;	Best Local Similarity	45.7%; Pred. No. 0.0017;
		Matches 223; Conservative	0; Mismatches 265; Indels 0; Gaps 0;
QY	98 TAATAAAAGATTCACTATCTGGGAATCGAACCGGGGTCTGTACTATGGCAAGATACTAT 157	QY	32 TCAATTAATTCAAAAAATATTTGTAATAAAAAACCGACCAAGTTGATCGGTATTTTA 91
DB	366 TCATAGCTCAATGACACGCGGGATCGAACCGGGTCTGTACCGTGGCAGGTACTAT 307	DB	5540 TAAATATACCATCAACAAAAATTCATATAAAAAATATCTAAAAAATTTAATAAATAT 5481
QY	158 TCTACCACTAGACATTTGGTTCATTTTGGTTTAAAGACTGCTTTTATTTGATTTACTC 217	QY	92 ATTATGTAATAAAAAAGATTCACTATCTGGGAATCGAACCGGGGTCTGTACTATGCAAGA 151
DB	306 TCTACCACTAGACATTTGGTTCATTTTGGTTTAAAGACTGCTTTTATTTGATTTACTC 247	DB	5480 CAAAAATAAAAAACTCTTCTAACTAAAAAACTATCAAAAAATACATTCAAAAATAAT 5421
QY	218 TTTAATTAATTTTTCACGAAATATACCGCAAAAGTTGGTCTGATTTTATTAATAAGTA 277	QY	152 TACTATTTACCACTAGACATTTGGTTCATTTTGGTTTAAAGACTGCTTTTATTTGATTT 211
DB	246 ATTATAAAGTATCATGTAAACAAACACACAATCTTTTGTAGTTTGTACCACTACGAAGTA 187	DB	5420 ATTTAAATAAACCTTAAAAATTAATAATATTTTATTTCAACATTAATAAAAAAATAAT 5361
RESULT 13		QY	212 ATACTCTTTTAAATTAATTTTTCACGAAATATACCGCAAAAGTTGGTCTGATTTTATTA 271
ABL32085/c		DB	5360 AAATATTTAAAAAACAATTTTCAACTAACCAACCCCAATCCCAACAAACAAAAATAAAA 5301
ID	ABL32085 standard; DNA; 5945 BP.	QY	272 AAAGTAAAAATTACTTTACCAAAAGTTGGTCTGATTTTAAAAATGATCCGCGGAATTAACCGA 331
XX		DB	5300 AAATCTCAAAAACAATAAAAAATAAAAAAATAATTTTCTCACACCCACACGTTATACTCA 5241
AC	ABL32085;	QY	332 CCAATTTTGGTAGGTTTTTAAATTAATTAATTTTAAATTTTAAATTTTAAATTTGAAACAA 391
XX		DB	5240 ACTACTCTAAAAAATAATATATATCTAATATTTTAACTAACCACTACATCATTAATAATAAT 5181
DT	26-MAR-2002 (first entry)	QY	392 AAGTTAGTCGGTTTCTTGAACACATAAAATTCGCGGGACTCAAAATAAGTTTCCCGCATTT 451
XX	Human immune system associated gene SEQ ID NO: 58.	DB	5180 AATTACTATCATTTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5121
XX	Human; immune system disease; cytosine methylation; antiasthmatic;	QY	452 TTGCGCCCAAGAAAAACCGACCAAGTTGGTTCGGTTTCGTAAAAAATAAAAAAATAAAAA 511
KW	antiarteriosclerotic; antianemic; cytosine methylation; antiasthmatic;	DB	5120 ATATATCAAAATAAACTTAAAAAATAAATCATTTTTCATATATATAATAAAAAAATAAAAA 5061
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;	QY	512 AATATATT 519
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;	DB	5060 ATCAAAAT 5053
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;		
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;		
OS	Homo sapiens.		
XX			
XX	WO200200928-A2.		
PN			
XX			
PD	03-JAN-2002.		
XX			
XX	02-JUL-2001; 2001WO-EP007537.		
XX			
XX	30-JUN-2000; 2000DE-01032529.		
XX			
XX	01-SEP-2000; 2000DE-01043826.		
XX			
XX	(EPIG-) EPIGENOMICS AG.		
XX			
XX	Olek A, Piepenbrock C, Berlin K;		
XX			
XX	WPI; 2002-130909/17.		
XX			
XX	Nucleic acid comprising fragment of chemically modified gene, useful for		
PT	diagnosis and treatment of diseases associated with abnormal cytosine		
PT	methylation.		
XX			
XX	Claim 1; SEQ ID NO 58; 32pp + Sequence Listing; German.		
PS			
XX			
CC	The present invention provides a number of human immune system associated		
CC	genes which are modified by the methylation of cytosines. The sequences		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		

PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 XX Claim 1; SEQ ID NO 2145; 32pp + Sequence Listing; German.
 PS The present invention provides a number of human immune system associated
 XX genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 XX Sequence 15518 BP; 3047 A; 362 C; 4183 G; 7926 T; 0 U; 0 Other;
 Query Match 7.0%; Score 62.8; DB 6; Length 15518;
 Best Local Similarity 46.6%; Pred. NO. 0.0032;
 Matches 234; Conservative 0; Mismatches 267; Indels 1; Gaps 1;
 QY 34 AATTAAATTCAAAAAATATTGTAAAAAACCACCAAGTTGATCGGTATTTAAT 93
 Db 942 AACTACGAACCTAAAAATAAAAAACAACTTCAAAATCTATTATAAAATTTCTTT 883
 QY 94 TATGTATAAAGAATTCATCTGGGAATCGAACCGGGGTCTGTACTATGCGAAGATA 153
 Db 882 ACTAAAAAATAAAAAAATTTCTAAACAATCAACACATATTAAATTTAATAAATA 823
 QY 154 CTATCTACCACTAGACATTTGTTTAAAGTCTGTTTAAAGTCTGTTTAAATTTAT 213
 Db 822 AAAAATACTATTAAAAATTAATACTAATACTCTCCATCTATTCTATATTTCTTAAAT 763
 QY 214 ACTCTTTAATTAATTTTGCACGAAATAACCGCAAAAGTTGTCGATTTTATTAATA 273
 Db 762 AATTATATCTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAAATAACTACACATA-AA 704
 QY 274 AGTAAATTTACTTACCAAACTATAATAACCAAAATCAACCACTCACTCACTCACT 333
 Db 643 CGGTAAATTTCTTTTATTAATTAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 584
 QY 394 GTTAGTCGGTTTCTTGAACATATAATTTTCGGGGACTCAAAATAGTTTCCCGCATTTT 453
 Db 583 CTCGCTACGTTACCAAACTATAATAACCAAAATCAACCACTCACTCACTCACTCACT 524
 QY 454 GCGCCAAAGAAACCGCAAAAGTTGTCGGTTTCTGTAATAAAAAAATAAAAAAATA 513
 Db 523 AACAAATATAATCTATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 464
 QY 514 TATATTTTAAAAACCGACCAA 535
 Db 463 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 442
 RESULT 15
 ABL34624/c
 ID ABL34624 standard; DNA; 15518 BP.
 XX
 AC ABL34624;
 XX
 DT 26-MAR-2002 (first entry)

XX Human metastasis associated gene SEQ ID NO: 177.
 DE Metastasis associated gene; cytostatic; gene therapy; cancer;
 XX cytosine methylation; gene; ds.
 KW Homo sapiens.
 OS
 XX WO200177376-A2.
 PN 18-OCT-2001.
 XX 06-APR-2001; 2001WO-EP003970.
 PF 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-010922/01.
 XX New nucleic acid derived from chemically treated metastasis genes, useful
 PT for diagnosis of cancers by analysis of cytosine methylation, also for
 PT treatment.
 XX Claim 1; SEQ ID NO 177; 23pp + Sequence Listing; English.
 PS The present invention provides a number of human metastasis associated
 XX genes which are modified by cytosine methylation. The sequences can be
 CC used in the diagnosis and treatment of cancer. The present sequence is
 CC one of the genes of the invention
 XX
 XX Sequence 15518 BP; 3047 A; 362 C; 4183 G; 7926 T; 0 U; 0 Other;
 Query Match 7.0%; Score 62.8; DB 6; Length 15518;
 Best Local Similarity 46.6%; Pred. NO. 0.0032;
 Matches 234; Conservative 0; Mismatches 267; Indels 1; Gaps 1;
 QY 34 AATTAAATTCAAAAAATATTGTAAAAAACCACCAAGTTGATCGGTATTTAAT 93
 Db 942 AACTACGAACCTAAAAATAAAAAACAACTTCAAAATCTATTATAAAATTTCTTT 883
 QY 94 TATGTATAAAGAATTCATCTGGGAATCGAACCGGGGTCTGTACTATGCGAAGATA 153
 Db 882 ACTAAAAAATAAAAAAATTTCTAAACAATCAACACATATTAAATTTAATAAATA 823
 QY 154 CTATCTACCACTAGACATTTGTTTAAAGTCTGTTTAAAGTCTGTTTAAATTTAT 213
 Db 822 AAAAATACTATTAAAAATTAATACTAATACTCTCCATCTATTCTATATTTCTTAAAT 763
 QY 214 ACTCTTTAATTAATTTTGCACGAAATAACCGCAAAAGTTGTCGATTTTATTAATA 273
 Db 762 AATTATATCTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAAATAACTACACATA-AA 704
 QY 274 AGTAAATTTACTTACCAAACTATAATAACCAAAATCAACCACTCACTCACTCACT 333
 Db 703 AAAAATCTTTCAAAAAATAAAAAATCTCTAAAAATTAATTAATTAATTAATTAAT 644
 QY 334 AATTGTCGGTTTCTTGAACATATAATTTTTCGGGGACTCAAAATAGTTTCCCGCATTTT 393
 Db 643 CGGTAAATTTCTTTTATTAATTAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 584
 QY 394 GTTAGTCGGTTTCTTGAACATATAATTTTCGGGGACTCAAAATAGTTTCCCGCATTTT 453
 Db 583 CTCGCTACGTTACCAAACTATAATAACCAAAATCAACCACTCACTCACTCACTCACT 524
 QY 454 GCGCCAAAGAAACCGCAAAAGTTGTCGGTTTCTGTAATAAAAAAATAAAAAAATA 513
 Db 523 AACAAATATAATCTATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 464

Search completed: November 11, 2004, 08:10:27
Job time : 448.342 secs

This Page Blank (uspto)

Result No.	Query No.	Score	Match	Length	DB	ID	Description	
1	1	75.4	8.4	1101	9	CNS0021J7	AL061936 Drosophil	
2	2	70.4	7.8	863	8	BZ088961	BZ088961 llasae08.	
C 3	3	67.2	7.5	1032	9	CNS020K7	AL175696 Tetraodon	
C 4	4	66.8	7.4	909	8	AZ670108	AZ670108 ENTMP51TF	
5	5	66	7.3	605	9	CL539617	CL539617 OS_Ba005	
6	6	65.8	7.3	759	9	CL799614	CL799614 OR_Cba001	
7	7	65.8	7.3	760	9	CL829972	CL829972 OR_CBa005	
C 8	8	65.8	7.3	1101	9	CNS00EVL	AL069706 Drosophil	
9	9	65.6	7.3	642	9	CG962071	CG962071 MBEDE79TR	
C 10	10	65.6	7.3	905	9	CNS00KHX	AL077798 Drosophil	
C 11	11	65.4	7.3	748	8	BZ413877	BZ413877 if18e02_g	
C 12	12	65.4	7.3	818	9	CG714001	CG714001 CGWDL03TV	
C 13	13	65.2	7.2	736	8	BZ755113	BZ755113 FUF8203TFV	
C 14	14	65.2	7.2	757	8	BZ991051	BZ991051 FUBNH65TD	
C 15	15	64.8	7.2	715	8	CG335052	CG335052 OGOA64TV	
C 16	16	64.8	7.2	1037	5	BQ648574	BQ648574 AGENCOURT	
C 17	17	64.6	7.2	802	9	CG730035	CG730035 CGWDR16TM	
C 18	18	64.6	7.2	919	9	CG442944	CG442944 QGT8J77TV	
C 19	19	64.4	7.1	836	9	CNS01100	AL099642 Drosophil	
C 20	20	64.4	7.1	845	9	CC935358	CC935358 ZUAH175TV	
C 21	21	64.4	7.1	856	6	CB624068	OS11Ea42E	
C 22	22	64.2	7.1	1029	9	CNS01ZGM	AL174271 Tetraodon	
C 23	23	64	7.1	651	9	CG954601	CG954601 PUFQX71TD	
C 24	24	64	7.1	653	9	CG954598	CG954598 PUFQX71TB	

BZ088961
 BZ088961.1 GI:23723478
 GSS.
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 863)
 Delehaanty,K., Fellw,G., Fulton,L., McCombie,W.R., Miner,T.,
 Nash,W., Rabinowicz,P.D. and Wilson,R.K.
 Whole genome shotgun reads from *Brassica oleracea*
 Unpublished (2002)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu

Plate: 0051 row: K column: 04
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES

Location/Qualifiers
1..605
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0051K04"
/tissue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OB_Ba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 7.3%; Score 66; DB 9; Length 605;
Best Local Similarity 73.7%; Pred. No. 0.0039;
Matches 84; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 100 ATAAAGAAATCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTATTC 159
Db 488 AAAAAGAAATCGACAGCCGGGAATCGAACCGGGTCTGTACTATGGCGAGGTACTATTC 547
QY 160 TACCCTAGACCATGTCTCATTTGTTTAAAGACTGTCTTTATTTGATTTAT 213
Db 548 TACCCTAGACCATGTCTCATTTGTTTAAAGACTGTCTTTATTTGATTTATTTGT 601

RESULT 6

CL799614 759 bp DNA linear GSS 06-AUG-2004
LOCUS OR_CBa0010001.f OR_CBa Oryza rufipogon genomic clone OR_CBa0010001
DEFINITION 5', genomic survey sequence.

ACCESSION

CL799614

VERSION

CL799614.1 GI:51023683

KEYWORDS

GSS.

SOURCE

Oryza rufipogon

ORGANISM

Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 759)
Kim, H., Yu, Y., Wissotski, M., Yost, D., Stum, D., Rao, K., Luo, M.,
Jetty, R., Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and
Wing, R.

TITLE

OMAP project

JOURNAL

Unpublished (2004)

COMMENT

Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0010 row: O column: 01
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES

Location/Qualifiers
1..759
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBa0010001"
/tissue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_CBa"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

ORIGIN

Query Match 7.3%; Score 65.8; DB 9; Length 759;
Best Local Similarity 66.7%; Pred. No. 0.0042;
Matches 94; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 104 AAAGATTCACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTATTC 163
Db 526 AAAAATGACCAAGCCGGGAATCGAACCGGGTCTGTACTATGGCGAGGTACTATTC 585
QY 164 ACTAGACCATGTCTCATTTGTTTAAAGACTGTCTTTATTTGATTTATTTAAT 223
Db 586 ACTAGACCATGTCTCTGTTGTAGTAGACAAATTTCTCAATTAATAGTCTATAT 645
QY 224 TATATTTTGCACGAATAA 244
Db 646 TTCCCTGCAAGCGCAACAA 666

RESULT 7

CL829972 760 bp DNA linear GSS 09-AUG-2004
LOCUS OR_CBa0051n11.f OR_CBa Oryza rufipogon genomic clone OR_CBa0051n11
DEFINITION 5', genomic survey sequence.

ACCESSION

CL829972

VERSION

CL829972.1 GI:51075582

KEYWORDS

GSS.

SOURCE

Oryza rufipogon

ORGANISM

Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 760)
Kim, H., Yu, Y., Wissotski, M., Yost, D., Stum, D., Rao, K., Luo, M.,
Jetty, R., Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and
Wing, R.

TITLE

OMAP project

JOURNAL

Unpublished (2004)

COMMENT

Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0051 row: N column: 11
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..760
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBa0051n11"
/tissue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_CBa"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

ORIGIN

Query Match 7.3%; Score 65.8; DB 9; Length 760;
Best Local Similarity 66.7%; Pred. No. 0.0042;
Matches 94; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 104 AAAGATTCACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTATTC 163
Db 526 AAAAATGACCAAGCCGGGAATCGAACCGGGTCTGTACTATGGCGAGGTACTATTC 585

```

QY 164 ACTAGACCATGGTTCATTTGTTTAAAGACTGCTTTATTGATTGATTACTCTTTAAT 223
Db 586 ACTAGACCATGGTTCGTTGCTGTTAGTACGACAAATTTCTCAATTAATAGTCTATAT 645

QY 224 TATATTTTTCACGAAATAA 244
Db 646 TTCCCTGCAAGCGCAACAA 666

RESULT 8
CNS00EVL/c 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly) genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
Genoscope.
Direct Submission
AUTHORS Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/clone_end="T7"

ORIGIN
Query Match 7.3%; Score 65.8; DB 9; Length 1101;
Best Local Similarity 32.4%; Pred. No. 0.0041;
Matches 188; Conservative 105; Mismatches 284; Indels 4; Gaps 1;

QY 21 ACTTTGTCGGTCAATTAATCAAAAAAATATTTGAAAAAACCACCAAGTTGA 80
Db 1018 AAWTTAATAATATATATWATWATATATATATATAWATAWATAWATAWATAWATAWAT 959

QY 81 TCGGTATTTTATGTATATAAAGATTCACATCTCTGGAATCGAACCGGGTCTGTA 140
Db 958 AWAATTTTAAATAAATAATWAAWTTATTTTATTTTATTTTATTTTATTTTATTTATW 899

QY 141 CTATGGCAAGATCACTATCTACCACTAGACCAATTTGGT---TCATTTTGGTTTAAAGACTG 196
Db 898 WTWAAAWAAWAAWAAWTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 839

QY 197 TCTTTTATTCATTTACTCTTTTAAATATATTTTTCACGAAATAACCGACCAAGTT 256
Db 838 TATWATWATATCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 779

```

```

QY 257 GGTGATTTTATTAATAAGTAAATTAATCTTACCAAGTTGGTTCGATTTTAAATGATC 316
Db 778 WATWTTTAAWAAWATATATATWATATTTATTTATTTATTTATTTATTTATTTATTT 719

QY 317 CGCCGAATTAACCGACCAATTTTGGTAGGTTTAAATATAAATTTTATTTATTTAA 376
Db 718 ATWATTTTATTTTATTTTATTTTATTTTAAATTTTAAATTAATWAAATWATTT 659

QY 377 TTGAAAACTAACCAAGTTAGTCGGTTTCTTGAACAATAAATTTTCGGCGGACTCAAAA 436
Db 658 ATTAAATTAATAATTAATAATTTTAAATAATTTTAAATAATTTTATTTATTTATTA 599

QY 437 TAGTTTCCCGCATTTTTCGCCAAAGAAACCGACCAAGTTGGTTCGTTTCGTAATAA 496
Db 598 WATTTTATAWATTAATATATATWATTTTAAATTTTATTTTAAATTTATTTTATTT 539

QY 497 AAAAAAATTTTAAATAATATTTTAAAAAAACCGACCAACTTTAGTCGGTTTTCGGTCG 556
Db 538 TTAAATAWAAWTTAAWAAATAATWAAWAAATAATWAAWAAATAATWAAWAAATAAT 479

QY 557 ATTTTTCACCGACCAAAAGTTGGTCGGTCGACCTTTGGTTCGG 597
Db 478 WWAATTTWKKKKKXANADTGAGARATKTDKKKKGGSGS 438

RESULT 9
CG962071 642 bp DNA linear GSS 15-DEC-2003
LOCUS MBEDE7STRC mth2 Medicago truncatula genomic clone 31M13, genomic
DEFINITION survey sequence.
ACCESSION CG962071
VERSION CG962071.1 GI:39884351
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 642)
AUTHORS Town,C.D., Shetty,J., Koo,H. and Feldblyum,T.F.
TITLE Sequencing of BAC ends from Medicago truncatula
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: MBEDE79TFC
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
Seq primer: CAGGAACAGCTATGACC
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..642
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="genotype Al7"
/db_xref="taxon:3880"
/clone="31M13"
/clone_lib="mth2"
/clone_end="T7"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Cook, D.R. and Kim, D.J, unpublished"

ORIGIN
Query Match 7.3%; Score 65.6; DB 9; Length 642;
Best Local Similarity 76.9%; Pred. No. 0.0047;
Matches 80; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 99 AATAAAGATTCACATCTCTGGAATCGAACCGGGTCTGTACTATGGCAAGTACTATT 158
Db 535 AAAAAAATTTGCACCGCGGAAATCGAACCGGGTCTGTACTATGGCAAGTACTATT 594

```



```

Db      330  CACTGGTGCAATTTGTCCTAAGTTCTTAATAAGGTTGATTAGGACCGTGCTACATGT 271
QY      231  TTGCA 235
Db      270  TCACA 266

RESULT 12
LOCUS   CC714001/c
DEFINITION CC714001 818 bp DNA linear GSS 19-JUN-2003
           OCWDL09TV ZM 0.7-1.5_KB Zea mays genomic clone ZMMBma0547A18,
           genomic survey sequence.
ACCESSION CC714001
VERSION   CC714001.1 GI:32118777
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 818)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
          Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE     Consortium for Maize Genomics
JOURNAL   Unpublished (2002)
COMMENT   Other GSSs: OGWDL09TH
          Contact: Cathy Whitelaw
TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..818
     organism="Zea mays"
     mol_type="genomic DNA"
     strain="B73"
     db_xref="taxon:4577"
     clone="ZMMBma0547A18"
     /clone_lib="ZM 0.7-1.5_KB"
     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
     methylation filtered genomic DNA library"

ORIGIN
Query Match      7.3%; Score 65.4; DB 9; Length 818;
Best Local Similarity 63.2%; Pred. No. 0.0051;
Matches 117; Conservative 0; Mismatches 66; Indels 2; Gaps 1;

QY      53  ATTGTAAAAAACCACCAAGTTGATCGGTATTTTAATTATGTAATAAAAGA--TT 110
Db      489  ATTAAACACATGATAGAAAAAATAATGGGAAAAAGAAATGAATAAAAAAATTG 430
QY      111  CACTATCTGGAAATCGAACCGGGTCTGTACTATGCGAAGATACTATTCTACCACTAGAC 170
Db      429  CACCAGCGGGAATCGAACCGGGTCTGTACCGTGGCAGGACTATTCTACCACTAGAC 370
QY      171  CATTGGTTCATTTTGTTTTAAGACTCTCTTTATTATTGATTATCTCTTAATTATTTT 230
Db      369  CATGTGTGCAATTTGTGCCTAAGTCTTAATAAGGTTTGATTATAGACCGTGCTACATGT 310
QY      231  TTGCA 235
Db      309  TCACA 305

RESULT 13
LOCUS   BZ755113/c
DEFINITION BZ755113 736 bp DNA linear GSS 10-MAR-2003
           PUFB203TD ZM 0.6-1.0_KB Zea mays genomic clone ZMMBta284B06,
           genomic survey sequence.
ACCESSION BZ755113
VERSION   BZ755113.1 GI:29232461
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 736)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA

```

```

ACCESSION BZ755113
VERSION   BZ755113.1 GI:28907462
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 736)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..736
     organism="Zea mays"
     mol_type="genomic DNA"
     strain="B73"
     db_xref="taxon:4577"
     clone="ZMMBta284B06"
     /clone_lib="ZM 0.6-1.0_KB"
     /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
     Cot selected genomic DNA library"

ORIGIN
Query Match      7.2%; Score 65.2; DB 8; Length 736;
Best Local Similarity 67.9%; Pred. No. 0.0056;
Matches 91; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY      99  AATAAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGCGAAGATACTATT 158
Db      403  AATTAAAAATTGCACCGCCGGGATCGAACCGGGTCTGTACTGCGCAGGACTATT 344
QY      159  CTACACATAGACCATTTGTTTCATTTTGTTTTAAGACTGTCTTTTATTGATTATCTCT 218
Db      343  CTACCACTAGACCACTGGTCTTTTGTGTTGTTCTTTTAAATTAGTTTAAATTATCTGT 284
QY      219  TTAATTATATTTT 232
Db      283  TCATCAATACATGT 270

RESULT 14
LOCUS   BZ991051/c
DEFINITION BZ991051 757 bp DNA linear GSS 25-MAR-2003
           PUFH65TD ZM 0.5-1.0_KB Zea mays genomic clone ZMMBta094K10,
           genomic survey sequence.
ACCESSION BZ991051
VERSION   BZ991051.1 GI:29232461
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 757)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA

```

Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES

Location/Qualifiers
 1..757
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBta094K10"
 /clone_lib="ZM 0.6-1.0 KB"
 /note="vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
 Cor selected genomic DNA library"

ORIGIN

Query Match 7.2%; Score 65.2; DB 8; Length 757;
 Best Local Similarity 67.9%; Pred. No. 0.0056;
 Matches 91; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
 Qy 99 AATAAAAGATTCTACTCTGGGAATCGAACCGGGGCTCTGACTATGCGCAAGATCTATT 158
 Db 267 AATTAAAAATTGCACCGCGGAATCGAACCGGGGCTCTGACTATGCGCGGCTACTATT 208
 Qy 159 CTACCACTAGACCACTGGTTTCATTTCTTTTAAGACTGCTCTTTTATTGATTATCTCT 218
 Db 207 CTACCACTAGACCACTGGTCTTTGTTGCTTTTAAATTAGTTTAAATTAATCTGT 148
 Qy 219 TTAATTATATTTT 232
 Db 147 TCATCAATACATGT 134

RESULT 15

CC335052 715 bp DNA linear GSS 16-MAY-2003
 LOCUS OGOAX64TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0341L08,
 DEFINITION genomic survey sequence.

ACCESSION CC335052
 VERSION CC335052.1 GI:30804465

KEYWORDS

GSS.

SOURCE

Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 715)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.

TITLE

Consortium for Maize Genomics

JOURNAL

Unpublished (2002)

COMMENT

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..715

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0341L08"

/clone_lib="ZM 0.7-1.5 KB"

/note="vector: pBSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 7.2%; Score 64.8; DB 8; Length 715;
 Best Local Similarity 66.4%; Pred. No. 0.0068;
 Matches 93; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 Qy 99 AATAAAAGATTCTACTCTGGGAATCGAACCGGGGCTCTGACTATGCGCAAGATCTATT 158
 Db 230 AATAAAAGATTCTACTCTGGGAATCGAACCGGGGCTCTGACTATGCGCGGCTACTATT 289
 Qy 159 CTACCACTAGACCACTGGTTTCATTTCTTTTAAGACTGCTCTTTTATTGATTATCTCT 218
 Db 290 CTACCACTAGACCACTGGTCTTCTGCTGATAATGATTTTCTGTTTCTTTCTTTCTCTTA 349
 Qy 219 TTAATTATATTTTGCACGA 238
 Db 350 ATTGACAGATTGGATGGA 369

Search completed: November 11, 2004, 12:05:39
 Job time : 2800.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 07:45:44 ; Search time 92.8475 Seconds
(without alignments)
6897.553 Million cell updates/sec

Title: US-10-009-966C-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tgcgcgaattaccagacca.....atgacttcattcgtattctc 901

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.6	5.5	3454	4	US-09-270-767-11909
2	48.6	5.4	640681	4	US-09-790-988-1
3	48.4	5.4	216	4	US-09-248-796A-7073
4	46.8	5.2	2435	3	US-09-306-593-1
5	45.8	5.1	660	1	US-07-991-867B-32
6	45.8	5.1	660	1	US-08-107-755A-32
7	45.8	5.1	660	2	US-08-544-332-32
8	45.8	5.1	660	4	US-09-370-861A-32
9	45.8	5.1	1511	1	US-07-991-867B-8
10	45.8	5.1	1511	2	US-08-107-755A-8
11	45.8	5.1	1511	2	US-08-544-332-8
12	45.8	5.1	1511	4	US-09-370-861A-8
13	45.8	5.1	4810	3	US-08-852-629-11
14	45.8	5.1	4838	3	US-08-852-629-15
15	45.8	5.1	11131	4	US-10-204-708-28
16	45.6	5.1	718	3	US-08-998-416-673
17	45	5.0	1099	4	US-09-270-767-11371
18	45	5.0	1141	4	US-09-806-708B-22
19	45	5.0	6040	4	US-10-204-708-69
20	43.8	4.9	1141	4	US-09-806-708B-22
21	43.8	4.9	29604	3	US-08-781-891-207
22	43.8	4.9	29604	4	US-09-618-166-207
23	43.6	4.8	13830	4	US-09-614-981-8
24	43	4.8	1055	4	US-09-806-708B-23
25	43	4.8	3348	4	US-09-976-594-363
26	42.6	4.7	1024	4	US-09-328-475C-81
27	42.6	4.7	3212	3	US-08-673-814-1

Sequence 1, Appli
Sequence 7, Appli
Sequence 13, Appli
Sequence 1, Appli
Sequence 93, Appli
Sequence 73, Appli
Sequence 71, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 43, Appli
Sequence 1, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-270-767-11909
; Sequence 11909, Application US/03270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11909
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11909

Query Match 5.5%; Score 49.6; DB 4; Length 3454;
Best Local Similarity 46.5%; Pred. No. 0.028;
Matches 160; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

Qy 34 ANTTAAATCAAAAAATATGTGTAATAAAAAACCGACCAAGTGTGCTATTTTAAAT 93
Db 492 AAAAAAATAAAAAACCGATTCATATAAAACAACTTCGTCAATTTTAAATTC 551
Qy 94 TATGTAATAAAAAAGATTCACTACTCTGGGAATCGAACCGGGGTCTGTACTATGGCAAGATA 153
Db 552 AGCATATATACATATATAATTTTCTTATGGTTTCATACACACCAAGTGTCTAATCG 611
Qy 154 CTAATCTACCTAGACCAATGGTTCATTTGTTTAAAGACTGCTTTTATTTGATTTAT 213
Db 612 CAAATGTATATAATAATACCTCTTATATATCGTATATCTTTGAAATTTGAAATTTAGCAGA 671
Qy 214 ACTCTTTAATATATTTTGGACGAAATAACCGACCAAGTGTGTCGATTTTATTTAAA 273
Db 672 ACCTTGAATGAACCTTTTGGTCTATATATATCTTATATCTTATGTTGTCATGCGTCGAC 731
Qy 274 AGTAAATTAATCAACCAAGTGTGTCGATTTTAAATGATCCGCGCAATTAACCGACC 333
Db 732 AATTTGACTGTATACCAAT 791
Qy 334 AATTTTGTAGTTTATTTTAAATATATATATATATATATATATATATATATATATAT 377
Db 792 ACTTTGAGTAAATTTGTTTACAAAGAAATTCATGTTTCTCAAT 835

RESULT 2
US-09-790-988-1/c
; Sequence 1, Application US/09790988

Patent No. 6632935
GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEKI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 5.4%; Score 48.6; DB 4; Length 640681;
Best Local Similarity 48.4%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 173 TTGGTTTCATTTTGGTTTAAAGACTGCTTTTATTTGATTACTCTTTAATTATATTTT 232
DB 35848 TAGATATATATATATTTAAATTTTATTCATTTGAAATTTGAACTCTGATTTCTTTT 35789

QY 233 GCAGAAAATACCGACCAAGTTGGTGCATTTTATTTAAAGTAAATTTACTTACCAA 292
DB 35788 ATATTTATTTTTCAAAAATACATGTTTAAATTAATATAGATTTTATTAATGAAAAATA 35729

QY 293 GTTGTGCTATTTTAAATCATCGCGAATTAACCGACCAATTTTGGTAGGTTTTTT 352
DB 35728 TATTTTCTTTTATCAATTAAGTATAAAAAATTTTATCAAAAAATATATTTT 35669

QY 353 AATATTAATTTTATTTTAAATTTGAAAACTAAACCAAGTTAGTCGGTTTCTTTGAAA 412
DB 35668 CAAAAATAATTTTTCATTTATATTTACAAAATATCTAAATAAATTTAATTTTACA 35609

QY 413 CATAAATTTCCGGGACTCAAAATAGTTTCCCGCATTT 451
DB 35608 TATAAATAAAAGGATAATAATGAATAAAACTCAATT 35570

RESULT 3
US-09-248-796A-7073/c
Sequence 7073, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 7073
LENGTH: 216
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-7073

Query Match 5.4%; Score 48.4; DB 4; Length 216;
Best Local Similarity 62.3%; Pred. No. 0.026; Indels 0; Gaps 0;
Matches 76; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 99 AATAAAGATTCATCTCTGGGAATCGAACCGGGGCTCTGCTACTATCGCAAGATACTATT 158
DB 99 AATAAAGATTCATCTCTGGGAATCGAACCGGGGCTCTGCTACTATCGCAAGATACTATT 158

DB 209 AAACAAAAATGCAAGAACCGGGAATCGAACCCGGGGCCCAACGATGCGCAACGTTGGATT 150
QY 159 CTACCACTAGACCATTTGGTTTCATTTTGTGTTTAAAGACTGCTCTTTTATTTGATTATATCTCT 218
DB 149 TTACCACTAAACCAATTTCTTGCAACCGCGCTATTGAAGTGCCTTTGTTGTAATTTTGGCCCC 90
QY 219 TT 220
DB 89 TT 88

RESULT 4
US-09-306-593-1/c
Sequence 1, Application US/09306593
Patent No. 6184018
GENERAL INFORMATION:
APPLICANT: Li, Xin-Liang
APPLICANT: Ljungdahl, Lars G.
APPLICANT: Chen, Huizhong
APPLICANT: Ximenes, Eduardo A.
TITLE OF INVENTION: Beta-glucosidase Coding Sequences and Protein from
TITLE OF INVENTION: Orpinomyces PC-2
FILE REFERENCE: 31-98us
CURRENT APPLICATION NUMBER: US/09/306,593
CURRENT FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/084,494
EARLIER FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2435
TYPE: DNA
ORGANISM: Orpinomyces sp. PC-2
FEATURE:
NAME/KEY: CDS
LOCATION: (39)..(2009)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (87)..(2009)
US-09-306-593-1

Query Match 5.2%; Score 46.8; DB 3; Length 2435;
Best Local Similarity 52.9%; Pred. No. 0.11; Indels 3; Gaps 2;
Matches 146; Conservative 0; Mismatches 127; Indels 3; Gaps 2;

QY 148 AAGATACTATTCTACCACTAGACCATTTGGTTTCATTTGTTTAAAGACTGCTTTTATTG 207
DB 2334 AACACAAAGCTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2275

QY 208 ATTTACTCTTTAATTTATTTTTCACGAAATTAACCGCAAGTTGTCGATTTTA 267
DB 2274 ATTTGTTTCTGCTTATTTATTTTCTTAAATAATATAAATCTTCTTTAGAAAT-TTTA 2216

QY 268 TTAATAAGTAAATTTACTTACCAAGTTGGTTCGATTTTAAATGATCCCGCAATTTAA 327
DB 2215 ATATAATCATTTATTTATTTTCAATTAATTTGATATTTATTTTATTTTATTAATGATTTCA 2156

QY 328 CCGACCAATTTGGTAGG--TTTTTTTAAATTAATTTTATTTTATTTTATTTTAAATGAAAAC 385
DB 2155 AAAAAAGATTTATTTCTATAAATTTATTAATTAATTAATTAATTTTATTTTCTACTTATAG 2096

QY 386 TAACCAAGTAGTCGGTTCTTTGAAACATAAATTT 421
DB 2095 CAAAAGAGTGAATAAATAATTCATTTATTAAT 2060

RESULT 5
US-07-991-867B-32
Sequence 32, Application US/07991867B
Patent No. 5476781
GENERAL INFORMATION:
APPLICANT: Noyer, Richard W.
APPLICANT: Hall, Richard L.


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UFI14.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-07-991-867B-8

Query Match 5.1%; Score 45.8; DB 1; Length 1511;
Best Local Similarity 44.4%; Pred No 0.17;
Matches 226; Conservative 0; Mismatches 282; Indels 1; Gaps 1;

Qy 9 AATTACCGACCAACTTTGGTCGGTCAATTAATTCAGAAAAAATATGTGTAACCAAAACC 68
Db 863 AAATTCGTGATATAATTTTAATAAATATTTTAAATATTTATAATTTAAAAAATAATAAT 922
Qy 69 GACCAAGTGTGTCGGTATTTTAATATGTAATAAAGAGATTCACTATCTGGGAATCGAA 128
Db 923 AAACAGAGATAATGTTATTAAATTAATATTAATTAATAAATAATAGTTAAATTTAGAAGAA 982
Qy 129 CCGGGGTCTGTACTTGGCAAGTACTATTTCTACCCTAGACCATTTGGTTCATTTTCTGTT 188
Db 983 GCATATAATATATTATGATAATA-ATATTTAATAATATTCAGAAAAATATTAAAGATT 1041
Qy 189 TAGACTGTCCTTTTATTTGATTTTATCTCTTTTAATTTATATTTTTCGACGAAAAATACCGA 248
Db 1042 TATATATTTCAAATTTAAATATTATTAATTTAAATTTTATAACAAAATTTAAAAAATAAA 1101
Qy 249 CCAAGTTGGTCGATTTTATTTAAAAAGTAAAAATTACTTACCAAGTTGGTCGATTTTTTTT 308
Db 1102 CATATTTAGATATATCTTATAACAAAAATAGCAATATAAGTATATTATCTACCAATTT 1161
Qy 309 AAATGATCCGCCGAATTAACCGCAATTTTGGTAGGTTTTTTTAAATATTAATTTTATTT 368
Db 1162 CTATAGAAATTTTAAAAATTGTAATCATGTAAATATAAATAGACTATAATTTTATTAAATAAT 1221
Qy 369 TATTTTAAITGAAAACTAACCGAATTTAGTCGGTTCCTTGAAACATAAATTTCCGCGGA 428
Db 1222 TAGTAAATTTTAAAAAATAATTAATATATCTTAAAAATAAATTTTGGTAACTTTAATAATGTT 1281
Qy 429 CTCAAAAATAGTTTTCCCGCATTTTTTGGCGCAAGAAAAACCGACCAAGTTTGGTCGGTTTC 488

```


Qy	249	CCAAAGTGGTCGATTTTATTA	AAAGTAAATTA	CTACTACAAAGTTGGTCGATTTTTT	308
Db	4423	CATATTTAGATATATCTTTA	CAACAAAATAGCAATAT	TAAGTAATTTATATACTACCA	4482
Qy	309	AAATGATCCGCGCAATTAAC	CGACCAATTTTG	TGTAAGTTTTTTTAAATTAATTTTAT	368
Db	4483	CTATAGAATTTTAAATTTG	GAATCATATATAAAT	GACTATAATTTTATTAATAAT	4542
Qy	369	TATTTTAAATTGAAAACATA	CCAAAGTTAGTCGGTTCT	TGAAACATAAATTTTCGCGGGA	428
Db	4543	TAGTAAATTTAAACAAAAT	TAATATATCTAAAAATAAAT	TTGGTAACTTTTAATAATGTTT	4602
Qy	429	CTCAAAAATAGTTTCCCGCA	TATTTTTCGCCAACAGAAAC	CGACCAAGTTGGTCGGTTTC	488
Db	4603	TTCTATATAGTATAGTTAG	TTAAATATGNAATCA	CAATTAAGATTAATAAATTA	4662
Qy	489	GTAAAAA	AAAAAATTTAAAAATATA	517	
Db	4663	TAGAAAAAATTAATTAAT	TTAAAAAATTA	4691	

QY 509 AAAAATATATTTTAAA 525
||| ||| ||| |||
Db 485 AAACCTAATAAAAAAA 469

Search completed: November 11, 2004, 12:09:06
Job time : 95.8475 secs

```

RESULT 15
US-10-204-708-28/c
; Sequence 28, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204.708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 28
; LENGTH: 11131
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-28

```

	Query Match	5.1%;	Score 45.8;	DB 4;	Length 11131;
	Best Local Similarity	48.6%;	Pred. No. 0.27;		
	Matches	125;	Conservative 0;	Mismatches 132;	Indels 0; Gaps 0;
Qy	269	TAAAAAGTAAAAATTACTTACCAAAAGTTGGTCAGATTTTTTTTAAATGATCGCGCGAATTAAAC	328		
Db	725	TAAAAACGAAAAATTCAAAAACCAACTAAACCAATATAATAACTAGCTCTCTACTAAAA	666		
Qy	329	CGACCAATTTTGGTAGGTTTTTTTAAATATATAATTTTTTATTATTTTAAATGAAAAACTAA	388		
Db	665	ATACAAAATTAACTAAAAAATAATAACGAACGCTCTATATCCCACTAATTAAAAAAACTAA	606		
Qy	389	CCAAAGTTATGTCGGTTCCTTGAAACATAAAATTCGCGGAGCTCAAAATAGTTTCCCGCA	448		
Db	605	AACAAAAAATCGCTTTAAACCTAAAAACGAAAAATCGTATATAACCGGAATTTATACCAC	546		
Qy	449	TTTTTGGCCCAAGAAAAACCGACCAAGTTGGTCGGTTTCGTAAAAAATAAAAAAATTTA	508		
Db	545	ATACTCACCCCTAACACAAAAACGAACCTCGGCTCTCAAAAAAATAAAAAAATAAA	486		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 10:27:15 ; Search time 482.239 Seconds
(without alignments)
10091.039 Million cell updates/sec

Title: US-10-009-966C-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tggctgaattaccgaccca.....atgacttcattcgtattctc 901

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3625171 seqs, 2700493622 residues

Total number of hits satisfying chosen parameters: 7250342

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	64.8	7.2	1140	15	US-10-190-312A-97
C 2	64.8	7.2	1140	15	US-10-190-312A-119
C 3	64.2	7.1	2131	16	US-10-424-599-109423
C 4	64	7.1	5945	15	US-10-311-455-58
C 5	62.8	7.0	15518	15	US-10-311-455-2145
C 6	62.8	7.0	15518	15	US-10-240-485-177
C 7	62.4	6.9	2000	9	US-09-887-576-129
C 8	58	6.4	3673778	15	US-10-312-841-1
C 9	57.2	6.3	3673778	15	US-10-312-841-1
C 10	56.2	6.2	552	17	US-10-021-323-93
C 11	55.4	6.1	6239	15	US-10-311-455-1749
C 12	55.4	6.1	6239	15	US-10-240-453-171

C 13	55.2	6.1	5407	15	US-10-311-455-2064
C 14	54.6	6.1	368	9	US-09-834-975-47
C 15	53	5.9	6092	16	US-10-221-613-36
C 16	52.8	5.9	8979	15	US-10-311-455-757
C 17	52.8	5.9	8979	16	US-10-321-613-137
C 18	52.6	5.8	520	17	US-10-021-323-7699
C 19	52.4	5.8	396	17	US-10-021-323-17801
C 20	52.4	5.8	419	9	US-09-960-352-11234
C 21	51	5.7	16914	16	US-10-221-613-214
C 22	50.4	5.6	5756	15	US-10-311-455-559
C 23	50.4	5.6	15732	14	US-10-239-676-96
C 24	50.4	5.6	15732	15	US-10-240-453-108
C 25	50.2	5.6	469	17	US-10-021-323-16830
C 26	50.2	5.6	500	10	US-09-814-353-5831
C 27	50.2	5.6	500	10	US-09-814-353-12112
C 28	50.2	5.6	520	17	US-10-021-323-7699
C 29	50	5.5	5997	15	US-10-311-455-1598
C 30	49.8	5.5	3673778	15	US-10-312-841-2
C 31	49.6	5.5	5520	15	US-10-311-455-1491
C 32	49.4	5.5	361	18	US-10-425-115-138029
C 33	49.2	5.5	499	17	US-10-021-323-10189
C 34	49	5.4	6963	15	US-10-311-455-952
C 35	49	5.4	15046	15	US-10-311-455-2074
C 36	48.8	5.4	296	10	US-09-814-353-5940
C 37	48.8	5.4	296	10	US-09-814-353-12220
C 38	48.6	5.4	5867	17	US-10-433-793-181
C 39	48.6	5.4	6052	16	US-10-221-613-352
C 40	48.6	5.4	640681	9	US-09-790-988-1
C 41	48.4	5.4	875	14	US-10-198-846-2872
C 42	48.4	5.4	17142	14	US-10-239-676-206
C 43	48.4	5.4	17142	15	US-10-311-455-2080
C 44	48.4	5.4	17142	15	US-10-240-453-304
C 45	48.2	5.3	5641	15	US-10-311-455-1369

ALIGNMENTS

RESULT 1

US-10-190-312A-97/c
; Sequence 97, Application US/10190312A
; Publication No. US2003019468A1
; GENERAL INFORMATION:
; APPLICANT: Chromagenics B.V.
; APPLICANT: Kruckeberg, Arthur L.
; TITLE OF INVENTION: DNA sequences comprising gene transcription regulatory qualities
; TITLE OF INVENTION: methods for detecting and using such DNA sequences
; FILE REFERENCE: 2183-4993.1
; CURRENT APPLICATION NUMBER: US/10/190,312A
; PRIOR FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: 60/303,199
; NUMBER OF SEQ ID NOS: 1079
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of STAR A13
US-10-190-312A-97

Query Match	7.2%	Score 64.8;	DB 15;	Length 1140;
Best Local Similarity	60.0%	Pred. No. 0.0011;		
Matches 108;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0;
QY	98	TAATAAAGATTCTACTATCTGGGATCGAACCGGGTCTGTACTATCGGAGATACAT	157	
DB	366	TCATAGCTCAATGACACCGCGGGAATCGAACCGGGTCTGTACTATCGGAGATACAT	307	
QY	158	TCATACACTAGACATTGGTTTCATTTTGTATTAAGACTGTCTTTTATTTGATTATCTC	217	

Db 306 TCTACCACTAGACCACTGGTCTCTTTTGGATTAAACACGACTAAATATATATATAGAT 247
QY 218 TTTAATTATATTTTTCACGAGAAATAACCGACCAAAAGTTCGTGATTTTATTAAAAAGTA 277
Db 246 ATTATAAGTATCATGTAAACACACACATCTTTTGTAGTTTGTACCACTACGAGTA 187

RESULT 2

US-10-190-312A-119/c
; Sequence 119, Application US/10190312A
; Publication No. US20030199468A1
; GENERAL INFORMATION:
; APPLICANT: Chromagenics B.V.
; APPLICANT: Otte, Arie P.
; APPLICANT: Kruckeberg, Arthur L.
; TITLE OF INVENTION: DNA sequences comprising gene transcription regulatory qualities
; TITLE OF INVENTION: methods for detecting and using such DNA sequences
; FILE REFERENCE: 2183-4993.1
; CURRENT APPLICATION NUMBER: US/10/190,312A
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: 60/303,199
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 1079
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 119
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of STAR A35
US-10-190-312A-119

Query Match 7.2%; Score 64.8; DB 15; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.0011; Mismatches 0; Gaps 0;
Matches 108; Conservative 0; Indels 0;
QY 98 TAATAAAAAAGTTCACATATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGTACTAT 157
Db 366 TCATAGCTCAATGCACCGCGGGAATCGAACCGGGTCTGTACCGTGGCAGGTACTAT 307
QY 158 TCTACCACTAGACCACTGGTCTCTTTTGGATTAAACACGACTAAATATATATAGAT 247
Db 306 TCTACCACTAGACCACTGGTCTCTTTTGGATTAAACACGACTAAATATATATAGAT 247
QY 218 TTTAATTATATTTTTCACGAGAAATAACCGACCAAAAGTTCGTGATTTTATTAAAAAGTA 277
Db 246 ATTATAAGTATCATGTAAACACACACATCTTTTGTAGTTTGTACCACTACGAGTA 187

RESULT 3

US-10-424-599-109423
; Sequence 109423, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 109423
; LENGTH: 2131
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_69823C.1
US-10-424-599-109423

Query Match 7.1%; Score 64.2; DB 16; Length 2131;

Best Local Similarity 74.3%; Pred. No. 0.0018;
Matches 81; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 98 TAATAAAAAAGTTCACATATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGTACTAT 157
Db 1936 TTAAATTTAAATGCACACCGCGGGAATCGAACCGGGTCTGTACCGTGGCAGGTACTAT 1995
QY 158 TCTACCACTAGACCACTGGTCTCTTTTGGATTAAACACGACTAAATATATATAGAT 206
Db 1996 TCTACCACTAGACCACTGGTCTCTTTTGGATTAAACACGACTAAATATATATAGAT 2044

RESULT 4

US-10-311-455-58/c
; Sequence 58, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 58
; LENGTH: 5945
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-58

Query Match 7.1%; Score 64; DB 15; Length 5945;
Best Local Similarity 45.7%; Pred. No. 0.0029; Mismatches 265; Indels 0; Gaps 0;
Matches 223; Conservative 0; Indels 0;
QY 32 TCAATTTAAATTCACCAAAAAATATTGTAAAAAATCCATATAAAAAATATCTAAAAAATTTAATAAATAT 5481
Db 5540 TAATATACCATCAACACAAAAATTCATATAAAAAATATCTAAAAAATTTAATAAATAT 5481
QY 92 ATTATGTATATAAAGATTCATATCTGGGAATCGAACCGGGTCTGTACTATGGCAGA 151
Db 5480 CAAAAATATAAAAAATCTTCTTAATAAAAAATCAATCAAAAAATCAATCAAAAAATAT 5421
QY 152 TACTATCTTACCACCTAGACCACTGGTCTCTTTTGGATTAAACACGACTAAATATATATATAT 211
Db 5420 ATTATAAATACCTTAAAAATTAATATATATATATATATATATATATATATATATATATAT 5361
QY 212 ATACTCTTTAATATATATTTTGCACGAAAAATACCGACCAAAAGTTGGTCTGATTTATATA 271
Db 5360 AAATATTTAAAAACATTTTTCACCTAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5301
QY 272 AAAGTAAATATATCTTACCAAAAGTTGGTCTGATTTTAAATATGATCCGCGGAATTAACCGA 331
Db 5300 AAATCTCAAAACCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5241
QY 332 CCAATTTTGGTAGGTTTTTTTAAATATATATATATATATATATATATATATATATATATAT 391
Db 5240 ACTACTCTAAAAAATAAT 5181
QY 392 AGTTAGTCGGTTTCTTGAACATATAATTTTCGCGGACTCAAAATAGTTTCCCGCATTT 451
Db 5180 AATTACTATCATTTTAAATATATCTATAAATACCAACCAATTTTAAATAAATATATATATAT 5121
QY 452 TTGCGCCAAAGAAAAACCGACCAAAAGTTGGTCTGTTTCTGTAATAAAAAAATAAATAAATA 511

Db 5120 ATATATCAAAATAAACTTAAAAAATAATCAATTTTCATATATATATAAAAAATTAAA 5061

QY 512 AATATATT 519

Db 5060 ATCAAATT 5053

RESULT 5

```

US-10-311-455-2145/c
; Sequence 2145, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation of Cytosine
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2145
; LENGTH: 15518
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2145

```

Query Match	7.0%	Score	62.8	DB 15	Length	15518			
Best Local Similarity	46.6%	Pred. No.	0.0075						
Matches	234	Conservative	0	Mismatches	267	Indels	1	Gaps	1

Qy	34	AAATAAATTCAAAAAATAATGTAAAAAAGCGACCAAGTGTGATCGTATTTTAAT	93
Db	942	AACTACGAACCTTAAAAATAAAAAACAAAAAATCTTAATATAAAAAATTTCTTT	883
Qy	94	TATGTAAATAAAAAAGATTCACATATCTGGGAATCGAACCGGGGTCTGTACTATGCGAAGATA	153
Db	882	ACTAAAAAATAAAAAAATAATCTAAACAAATCAAAACACATATTAAAAATTTAATAAATA	823
Qy	154	CTATCTCACACTAGACCAATCGGTCATATTTGGTTTAAAGACTGTCCTTTATTTGATTTAT	213
Db	822	AAAAATAACATATTAAAAAATAATACTAATACTCTCCATACTATCTATATATCTTTAAAAA	763
Qy	214	ACTCTTTAATATATATTTTTTGCACGAAAAATAACCGACCAAGTGGTCGATTTTATTTAAAAA	273
Db	762	AAATTAATCTTTTTTTTTTTTTTAAAAAAACTAAAAAATAAATCAACAATA-AACAA	704
Qy	274	AGTAAATAATPACTTACCAAGTGGTCGATTTTTTTAAATGATCGCGCAATTAACCGACC	333
Db	703	AAAAATCTTTCAAAAAATAAAAAATACTCTAAAAATTAATTAATAATAATTAACAAAT	644
Qy	334	AAATTTGGTAGCTTTTTTATATTAATTTTTTATTTATTTAAITGAAAACTTAACCAA	393
Db	643	CGGTAAATTTCTTTTATATCTATATTTCTTTCTTTTCTTTTAAATAAAACTAAAT	584
Qy	394	GTTAGTCGGTTTTCTTGAACACATAAATTTCCGGGGACTCAAAAAATAGTTTCCCGCATTTTT	453
Db	583	CTCGCTACGTTTACCAAAACTATAATAAACCAAAATCACACCCTACACTCCAACCTAAAT	524
Qy	454	GGCCCAAGAAACCGACCAAGTTGGTCGGTTTGGTAAAAAATAAAAAAATTTAAAAAA	513
Db	523	AACAAATAAATTCCTATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	464
Qy	514	TATATTTTAAAAAACCGACCAA	535

Db 463 AAAAAAAAAAAAAAAAAAAAAA 442

RESULT 6

```

US-10-240-485-177/c
; Sequence 177, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 177
; LENGTH: 15518
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-177

```

Query Match	7.0%	Score 62.8;	DB 15;	Length 15518;
Best Local Similarity	46.6%;	Pred. No. 0.0075;		
Matches 234;	Conservative	0;	Mismatches 267;	Indels 1;
Gaps 1;				

[illegible]

```

Db      463  AAAAAAAAAAAAAAAAAAAAAA 442
          ||||| | |
RESULT 7
US-09-887-576-129
; Sequence 129, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-129

Query Match      6.9%; Score 62.4; DB 9; Length 2000;
Best Local Similarity 81.8%; Pred. No. 0.0042;
Matches 72; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      99  AATAAAAGATTTCACATCTCGGAATCGACCGGGGTCTCTACTATGCGCAAGATACATT 158
Db      1624 ACTAAAATTTGCACCGCGGGAATCGAACCGGGGTCTCTACTATGCGCAAGATACATT 158
QY      159  CTACCACTAGACCACTTGGTTCATTTTGT 186
Db      1684 CTACCACTAGACCACTGGTCTTCTTGT 1711

RESULT 8
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match      6.4%; Score 58; DB 15; Length 3673778;
Best Local Similarity 47.1%; Pred. No. 0.61;
Matches 178; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

```

Db 2573855 AAATAAAATATTAATACAAATTTAAATAATCTTTAAATTTAAATCTCTAAATTTTAT 2573796
QY 345 GTTTTTTAAATATTAATTTTATTTTAAATTTGAAATACTAAACCAAGTTAGTCGGTT 404
Db 2573795 TTTTATTTAAATATTTTAAATATATATATTTATATA-TATATATAATATATATAT 2573737
QY 405 TCTTGAAACATAATTTTCGGGAGCTCAAAATAGTTTCCCGCATTTTTCGCCCAAGAA 464
Db 2573736 ATACAATACAAAAATAAATTAATAAACAATTCCTATTCATCTCTCCCTCCCAATAAAA 2573677
QY 465 AACCGACCAAGTTGTCGGTTCGTGTAATAAAAAAATTTAAATAATATATTTTAA 524
Db 2573676 ATAAAAATCTTTACCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2573617
QY 525 AA 526
Db 2573616 TA 2573615

RESULT 10
US-10-021-323-93/c
; Sequence 93, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 93
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(552)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3825-001-Q1-K6-C5
US-10-021-323-93

Query Match 6.2%; Score 56.2; DB 17; Length 552;
Best Local Similarity 44.8%; Pred. No. 0.052;
Matches 202; Conservative 0; Mismatches 249; Indels 0; Gaps 0;
QY 85 TATTTTAAATATCTAATAAAGATTCACCTATCTGGGAATCGAACCGGGTCTGTACTAT 144
Db 470 TTTTATTTTATTTTNNAAAAATATTTTNNNANCACTAAACCTCCAAAAA 411
QY 145 GCGAAGATCTATCTACCACTAGACCATTTGTTTCATTTTGTGTTTAAAGACTGTCTTTAT 204
Db 410 ANNTTTTAAATTTTATTTTAACTCCCTCCCTTTTTCCTTTTCAAAAAATTTT 351
QY 205 TCGATTTACTCTTTAAATATATTTTTCGACGAAATACCGACCAAGTTGTCGGT 264
Db 350 TCTCTCCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 291
QY 265 TTATTTAAAGTAAATTTACTTACCAAGTTGTCGATTTTATTTTAAATGATCGCGCAAT 324
Db 290 TAAATTTTATTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 231
QY 325 TAAACCGCAATTTTGGTAGGTTTTTAAATATATTTTATTTTATTTTATTTTATTTTAA 384
Db 230 TTTTTCACCCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 171
QY 385 CTAAACCAAGTTAGTCGGTTCTTTGAAACATAAATTCGGGAGCTCAAAATAGTTTCC 444

Db 170 AAAATTTAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 111
QY 445 CGCATTTTTCGCCCAAGAAACCGACCAAGTTGTCGGTTTCGTAATAAAAAA 504
Db 110 ATTTTATTTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 51
QY 505 TTTTAAAAATATATTTTAAAAAATCGACCA 535
Db 50 AAAAAAATAAAAAAATAAAAAAATAAAAAA 20

RESULT 11
US-10-311-455-1749/c
; Sequence 1749, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PISPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determini
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1749
; LENGTH: 6239
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1749

Query Match 6.1%; Score 55.4; DB 15; Length 6239;
Best Local Similarity 46.9%; Pred. No. 0.19; Indels 0; Gaps 0;
Matches 173; Conservative 0; Mismatches 196;
QY 172 ATGGTTCATTTTGTGTTTAAAGACTGTCTTTTATTTGATTTTATCTTTTATTTATTTT 231
Db 1844 ATATATTTTATTTTATTTTAAATTAACGTTTATTAATTTTATTAATTTTATTTTCTT 1785
QY 232 TGCACGAAATPAACCGACCAAGTTGTCGATTTTATTAATAAGTAAATTAATTAATTA 291
Db 1784 TTATTAATATAATAAATAATACTTATTTATTTATTTATTTATTTATTTATTTATTTATA 1725
QY 292 AGTTGGTCGATTTTATTTTAAATGATCCGCGAATTAACCGCAATTTTGGTAGGTTTTT 351
Db 1724 TCTTATTTATTTTATTTTAAATTTAAATTTATTTTATTTTATTTTATTTATTTATTT 1665
QY 352 TAAATTTAAATTTTATTTTATTTTAAATTTGAAATACTAAACCAAGTTAGTCGGTTTCTTGA 411
Db 1664 TATAATTAACAATAATCATCATACACCTATAATCCCACTACAAAAAATACTTAAAC 1605
QY 412 ACATAAATTCGGGAGCTCAAAATAGTTTCCCGCATTTTTCGCCCAAGAAACCGAC 471
Db 1604 CCAAAAAATCGAAACTCAACAAACCAATAATATATCATCTATCTCAACCTCAACGACA 1545
QY 472 CAAGTTGGTCGGTTTCGTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 531
Db 1544 AATAAACCCTACTCTCTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1485
QY 532 CCAACTTTA 540
Db 1484 TTAICTTTA 1476

RESULT 12

US-10-240-453-171/c
; Sequence 171, Application US/10240453
; Publication No. US20030148326A1

; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription

; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; TITLE OF INVENTION: With DNA Transcription

; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453

; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973

; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8

; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8

; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350

; SEQ ID NO 171
; LENGTH: 6239
; TYPE: DNA

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-240-453-171
Query Match 6.1%; Score 55.4; DB 15; Length 6239;
Best Local Similarity 46.9%; Pred. No. 0.19; Indels 0; Gaps 0;
Matches 173; Conservative 0; Mismatches 196

Qy 172 ATTGGTTCATTGTTTAAAGACTGCTTTTATGATTATATCTCTTTAAATATATATTTT 231

Db 1844 ATATATTTTATTTTAAATTAACGGTTTATTAATTTTATTAATTTTATTTTATTTCT 1785

Qy 232 TGCAGAAATACCGACCAAGTTGGTCGATTTTATTAATAAGTAATTTACTTACCAA 291

Db 1784 TTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1725

Qy 292 AGTGGTTCGATTTTAAATGATCCCGCAATTAACCGCAATTTTGGTAGGTTTTT 351

Db 1724 TCTATTATTTTAAATTTAAATAATTTCTTTTATTTATTTATTTATTTATTTATTT 1665

Qy 352 TAATATTAATTTTATTTATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 411

Db 1664 TATAATTAACAATAATCATCATACCACTATAATCCCACTACCAAAAAAATTTACTTAAAC 1605

Qy 412 ACATTAATTTTCGGGGACTCAAAATAGTTTCCGCAATTTTTCGCAAAAGAAACCGAC 471

Db 1604 CCAAAAAATCGAACTACAAACCAATAATAATATATCATCTATCTCAACCTTAACGACA 1545

Qy 472 CAAAGTTGGTGGTTTCGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 531

Db 1544 AAATAAACCTTCTCTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1485

Qy 532 CCAACTTTA 540

Db 1484 TTATCTTTA 1476

RESULT 13
US-10-311-455-2064/c

; Sequence 2064, Application US/10311455
; Publication No. US20030143606A1

; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation

; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455

; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 2064
; LENGTH: 5407
; TYPE: DNA

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-2064
Query Match 6.1%; Score 55.2; DB 15; Length 5407;
Best Local Similarity 47.6%; Pred. No. 0.2; Indels 0; Gaps 0;
Matches 162; Conservative 0; Mismatches 178

Qy 186 TTTTAAGACTGCTTTTATGATTATATCTCTTTAAATATATTTTTCACGAAATAAC 245

Db 2083 TATTATAATATCTCTATACCTTACCACCTCTTTTAAATTTTAAAAATAAAAAAC 2024

Qy 246 CGACCAAGTTGGTCGATTTTATTAATAAGTAATTTACTTACCAAGTTGGTCGATTTT 305

Db 2023 ATTTCAATTTTAAATTTATAACAATTTCAACTTTAAAACTTTACAAAACTCCTTATTT 1964

Qy 306 TTTAAATGATCCGCGAATTAACCGCAATTTTGGTAGGTTTTTAAATTAATTTT 365

Db 1963 AATTAATAATTAACCAATTTATCTACCACTTTTCTTAACTCTATTTTAAATTTCTAC 1904

Qy 366 ATTTATTTTAAATGAAAACTTAACCAAGTTAGTCGGTTTCTTGAACAATAATTTTCGCG 425

Db 1903 CCAAAATAAAAAAGCTATAAACAACAAAAATAATAATTTTATATCTATATAACAACCAATAT 1844

Qy 426 GGAATCAAAATAGTTTCCCGCATTTTTCGCAAAAGAAACCGCAAAAGTTGGTCGGT 485

Db 1843 TAAATCTCTACTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1784

Qy 486 TTCTGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 525

Db 1783 TAAATCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1744

RESULT 14
US-09-834-975-47/c

; Sequence 47, Application US/09834975
; Patent No. US20020110815A1

; GENERAL INFORMATION:
; APPLICANT: Brown, James
; APPLICANT: Bolt, Jeffrey

; APPLICANT: Van Huffel, Christophe
; APPLICANT: Van Huffel, Christophe

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975

; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538

; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 368
; TYPE: DNA

; ORGANISM: Homo sapiens

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(368)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-47

Query Match      6.1%; Score 54.6; DB 9; Length 368;
Best Local Similarity 44.2%; Pred. No. 0.096;
Matches 138; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 224 TATATTTTTCACGAAATAAACCGACCAAGATTGGTCGATTTTATTAATAAAGTAAATTA 283
Db      |||||
QY 355 TTTTATTTTAAATAAATAAACCCNCCCTTTTNTTTTAAATNAATAAATTAATTT 296
Db      |||||
QY 284 CTACCAAGTGGTCGATTTTAAATGATCCGCCGAATTAACCGACCAATTTTGGTA 343
Db      |||||
QY 295 TTCCCNCTTCCNCCNNAATTTTAAATGNGTTTTTTTTTTCNCCNACCATTTT 236
Db      |||||
QY 344 GGTATTTTAAATTAATTTTATTTTATTTTAAATGAAATTAACCAAGTTAGTCGGT 403
Db      |||||
QY 235 TTTTATTTTAAANNNGTTTAAATTTTNAATAAATAAATAAATTTTNAATAAATAAATTT 176
Db      |||||
QY 404 TTCTGAAACATAAATTCGCGGAGCTCAAAATAGTTTCCGCAATTTTGGCCCAAGA 463
Db      |||||
QY 175 TTTTATTTTAAATCAATANNNTTNAATTTTNAATTTTNNNTCCNNAATAAATAAATAAATAA 116
Db      |||||
QY 464 AACCACCAAGTGGTCGGTTTCGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 523
Db      |||||
QY 115 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 56
Db      |||||
QY 524 AATAACCGACCAA 535
Db      |||||
QY 55 AGNAATAAATAA 44
Db      |||||
```

```
RESULT 15
US-10-221-613-36
; Sequence 36, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 36
; LENGTH: 6092
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-36
```

```
Query Match      5.9%; Score 53; DB 16; Length 6092;
Best Local Similarity 51.0%; Pred. No. 0.6;
Matches 125; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 143 ATGCAAGATCTATTCTACCACTAGACCATGGTTCATTTTCTTTTAAGACTGCTCTTT 202
```

```
Db 2073 ATAATATGATTTTATTAATAAATAAATTTTAAATATACGTGTTTAAATAAATTTTA 2132
QY 203 ATTTGATTTTACTCTTTTAAATTAATTTTTCACGAAATAAACCGACCAAGTTGGTCGA 262
Db 2133 TATTTATATATAAATAAATAAAGTAATATAAAGTTTGTATATAAATTAATAAGTATAAGTA 2192
QY 263 TTTTATTAATAAAGTAAATAATTAATTTACCAAGTTGGTCGATTTTAAATGATCCGCCGA 322
Db 2193 TATTATTAAAGAAAGTATTGTTTAAAAAATTCGTGAGTTTGTGTTTTTTTGAATAATAA 2252
QY 323 ATTAACCGACCAATTTGGTAGGTTTAAATAATTAATTTTAAATTTTAAATTTAAATTTGAAA 382
Db 2253 GTTTATTTTATTATTGATGTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTGAAA 2312
QY 383 AACTA 387
Db 2313 AATA 2317
```

Search completed: November 11, 2004, 14:31:33
Job time : 493.239 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 06:56:19 ; Search time 4307.14 Seconds
(without alignments)
10990.348 Million cell updates/sec

Title: us-10-009-966c-2_COPY_2155_3155
Perfect score: 1001
Sequence: 1 aacatagacttttgattaaa.....accactactagaatccgg 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenEmbl.*													
	1:	gb_ba.*												
	2:	gb_btgi.*												
	3:	gb_in.*												
	4:	gb_on.*												
	5:	gb_ov.*												
	6:	gb_pat.*												
	7:	gb_ph.*												
	8:	gb_pl.*												
	9:	gb_pr.*												
	10:	gb_ro.*												
	11:	gb_sts.*												
	12:	gb_sy.*												
	13:	gb_un.*												
	14:	gb_vi.*												

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	DB ID	Description		
1	1000.2	99.9	4312	6	AX068013	Sequence	
2	997	99.6	3294	6	AX068012	Sequence	
3	997	99.6	4135	6	AX068019	Sequence	
4	980.6	98.0	3293	8	AF376772S1	Nicotiana	
5	100	10.0	3663	8	AB012639	Nicotiana	
6	66.6	6.7	154256	2	CR678065	Danio rer	
7	85.6	6.6	79018	9	AL627107	Human DNA	
8	64.6	6.5	207772	2	EX897691	Danio rer	
9	64.2	6.4	201329	2	CR450778	Danio rer	
10	64	6.4	156463	2	CR388040	Danio rer	
11	64	6.4	160534	2	CR450837	Danio rer	
12	63.6	6.4	201312	2	AC067900	Homo sapi	
13	63	6.3	154675	2	CR354440	Danio rer	
14	62.8	6.3	4753	8	NSY252274	Nicotiana	
15	62.6	6.3	90550	9	AL592166	Human DNA	
16	62.6	6.3	174001	2	EX927081	Danio rer	
17	62.6	6.3	193764	2	CR450831	Danio rer	
18	62.4	6.2	172811	9	AC066608	Homo sapi	
19	62.4	6.2	236582	2	CR388102	Danio rer	

20	62.4	6.2	349418	3	CR382398	Plasmodiu
21	61.8	6.2	185672	2	BX640474	Danio rer
22	61.4	6.1	177888	2	EX957314	Danio rer
23	61.2	6.1	245072	2	CR381551	Danio rer
24	61.2	6.1	348034	3	CR382400	Plasmodiu
25	61	6.1	163533	2	CR627488	Danio rer
26	61	6.1	190128	2	BX908801	Danio rer
27	61	6.1	233658	2	CR354421	Danio rer
28	60.8	6.1	186313	2	CR631129	Danio rer
29	60.6	6.1	181539	2	EX957351	Danio rer
30	60.4	6.0	1141	6	AX083744	Sequence
31	60.4	6.0	125836	2	AC137629	Homo sapi
32	60.4	6.0	156978	2	CR339045	Danio rer
33	60.2	6.0	110000	3	PFMAL1P2_1	Continuation (2 of
34	60.2	6.0	132112	2	CR628862	Danio rer
35	60	6.0	138982	2	CR388017	Danio rer
36	60	6.0	196756	2	CR556696	Danio rer
37	59.8	6.0	72243	9	AL731858	Human DNA
38	59.8	6.0	155156	2	CR627499	Danio rer
39	59.8	6.0	162206	9	AC108706	Danio rer
40	59.8	6.0	177423	2	CR356242	Danio rer
41	59.8	6.0	250663	3	AE014826	Plasmodiu
42	59.8	6.0	313050	3	PFAG29352	Plasmodiu
43	59.6	6.0	189963	2	CR559944	Danio rer
44	59.4	5.9	198340	2	CR556709	Danio rer
45	59.2	5.9	148856	2	BX942828	Danio rer

ALIGNMENTS

RESULT 1
AX068013
LOCUS AX068013 4312 bp DNA linear PAT 19-JAN-2001
DEFINITION Sequence 2 from Patent WO0077187.
ACCESSION AX068013
VERSION AX068013.1 GI:12329821
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 0077187-A 2 21-DEC-2000.
Roitsch, Thomas, Dr. - 1997;
FEATURES
source location/Qualifiers
1..4312
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN

Query Match	99.9%;	Score	1000.2;	DB	6;	Length	4312;	
Best Local Similarity	100.0%;	Pred. No.	7.6e-172;					
Matches 1001;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	AAATAGACTTTTGGATTAAATTAATTAATTTGTATGACAAAGGATTAAACCTAGTTAAT	60					
Db	2155	AAATAGACTTTTGGATTAAATTAATTAATTTGTATGACAAAGGATTAAACCTAGTTAAT	2214					
Qy	61	GATAATAACTTCAGATCTATATCAATTAACAGCAATCACCGTCTACAGCGTTGAGAG	120					
Db	2215	GATAATAACTTCAGATCTATATCAATTAACAGCAATCACCGTCTACAGCGTTGAGAG	2274					
Qy	121	AGATTAAATGTGATGTCATTCATATTTTCAAGATCATTTAATCATAGGGGAATATCAAG	180					
Db	2275	AGATTAAATGTGATGTCATTCATATTTTCAAGATCATTTAATCATAGGGGAATATCAAG	2334					
Qy	181	CAATAAATAACCAATAAATGGCATTAAAGTAATAAGGAGATGATTCACCCAAATATTGAA	240					

in start part
parent

ORIGIN

Query Match	98.0%;	Score 980.6;	DB 8;	Length 3293;
Best Local Similarity	99.3%;	Pred. No. 2.9e-168;		
Matches 994;	Conservative 1;	Mismatches 5;	Indels 1;	Gaps 1;
QY	1	AATATAGAGCTTTTGATTAAATTAATTAATTCATGATGACAAAGGATTAACCTAGTTAAT	60	
DB	1138	AATAATAGACTTTTGATTAAATTAATTAATTCATGATGACAAAGGATTAACCTAGTTAAT	1197	
QY	61	GATATATACCTTCAGATCTATATATCAATATACACGAATCAAGTCTATGACAGCGTTGAG	120	
DB	1198	GATAATAACTTCAGATCTATATATCAATATACACGAATCAC-GTATATAGCAGCGTTGAG	1256	
QY	121	AAGATTAAATGTGATGTGYCATTTCAATATTTCAAGATCATTAATGATAGGGAATATCAAG	180	
DB	1257	AAGATTAAATGTGATGTACATTCATATTTCAAGATCATTAATGATAGGGAATATCAAG	1316	
QY	181	CAATTAATTAACGATTAATATGCGATTAAAGTAATTAAGGAGAATGATTCACCCAAATTGAA	240	
DB	1317	CAATTAATTAACGATTAATATGCGATTAAAGTAATTAAGGAGAATGATTCACCCAAATTGAA	1376	
QY	241	TGAGGTGATGATTTCTCTTTTGTGACAAATGATGAATGATGGGCAATACTAGAATGTTGG	300	
DB	1377	TGAGGTGATGATTTCTCTTTTGTGACAAATGATGAATGATAGCAAAATACTAGAATGTTGG	1436	
QY	301	GACCCCTTCGGATCTAATGAATAAGATGGAATAGTAGATAATCGAATCTCTTTTAGAA	360	
DB	1437	GACCCCTTCGGATCTAATGAATAAGATGGAATAGTAGATAATCGAATCTCTTTTAGAA	1496	
QY	361	AGGTAGTGATTTGCTTTTATCTAGAGAGAAAGTCGCTTTTCAAAAGAATATTTTTATCAG	420	
DB	1497	AGGTAGTGATTTGCTTTTATCTAGAGAGAAAGTCGCTTTTCAAAAGAATATTTTTATCAG	1556	
QY	421	AGAAATTTACATCCGCTCTCCCTATCTCTTTTCTATTTATATGGGACATTCCTCAA	480	
DB	1557	AGAAATTTACATCCGCTCTCCCTATCTCTTTTCTATTTATATGGGACATTCCTCAA	1616	
QY	481	TCAATCTCTAAAGTACATACACAAAGAATATCAATAAAATATTTTTTGAATATTCAT	540	
DB	1617	TCAATCTCTAAAGTACATACACAAAGAATATCAATAAAATATTTTTTGAATATTCAT	1676	
QY	541	TATAAARACTAGCTGTAGACATCGACCTCGGTCGTTATGTACTACTCGGTACGAGCC	600	
DB	1677	TATAAARACTAGCTGTAGACATCGACCTCGGTCGTTATGTACTACTCGGTACGAGCC	1736	
QY	601	TGTCATTTACTAATCGACCTCGATTACATCCTTTCTACGATACGTCTCATGTCAAATC	660	
DB	1737	TGTCATTTACTAATCGACCTCGATTACATCCTTTCTACGATACGTCTCATGTCAAATC	1796	
QY	661	TTAATGAAGCAGATTTTGACCCATCAATATATGACAAATATGCTTCCAAAGAAACA	720	
DB	1797	TTAATGAAGCAGATTTTGACCCATCAATATATGACAAATATGCTTCCAAAGAAACA	1856	
QY	721	TGGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAAAGATCTGAATTTATTTAAGA	780	
DB	1857	TGGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAAAGATCTGAATTTATTTAAGA	1916	
QY	781	ATAGTGTTTTTCTTTTCTTTTTCATATCTAAGGATTAAGCAACCATGATAGAAAGG	840	
DB	1917	ATAGTGTTTTTCTTTTCTTTTTCATATCTAAGGATTAAGCAACCATGATAGAAAGG	1976	
QY	841	CTTAGTAACCTATATATCAAAAGGAATGGTGTGTTTTCTTTTAAATATGATAAAAAATTTGTG	900	
DB	1977	CTTAGTAACCTATATATCAAAAGGAATGGTGTGTTTTCTTTTAAATATGATAAAAAATTTGTG	2036	
QY	901	AATATAGAGATTAAGATCAATTAACAAAGGTTATGGTGGATGGAACGACGAGCGACC	960	
DB	2037	AATATAGAGATTAAGATCAATTAACAAAGGTTATGGTGGATGGAACGAGCGAGACC	2096	
QY	961	TATGTGTTATAGTAAGGGGTCACCCCACTACTAGAAATCCGG	1001	
DB	2097	TATGTGTTATAGTAAGGGGTCACCCCACTACTAGAAATCCGG	2137	

```

RESULT 5
AB012639/c
LOCUS
DEFINITION
AB012639
ACCESSION
AB012639.1
VERSION
GI:3036950
KEYWORDS
light harvesting chlorophyll a/b-binding protein.
SOURCE
Nicotiana sylvestris (wood tobacco)
ORGANISM
Nicotiana sylvestris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1
REFERENCE
1
AUTHORS
Hasegawa,K., Yukawa,Y., Sugita,M. and Sugiura,M.
TITLE
Organization and transcription of the gene family encoding
chlorophyll a/b-binding proteins in Nicotiana sylvestris
JOURNAL
Gene 289 (1-2), 161-168 (2002)
MEDLINE
22032967
PubMed
12036594
REFERENCE
2
BASES
1 to 3663
AUTHORS
Hasegawa,K., Yukawa,Y., Sugita,M. and Sugiura,M.
TITLE
Direct Submission
JOURNAL
Submitted (27-MAR-1998) Masahiro Sugiura, Nagoya University, Center
for Gene Research; Chikusa-ku, Nagoya, Aichi 464-8602, Japan
ID-mail:rh4979@nuc.cc.nagoya-u.ac.jp, tel:+81-52-789-3081,
Fax:+81-52-789-3081
FEATURES
Location/Qualifiers
1..3663
/organism="Nicotiana sylvestris"
/mol_type="genomic DNA"
/db_xref="taxon:4096"
/clone="lambda NSC81"
/tissue_type="leaf"
/clone_lib="lambda DASH genomic library"
2242..2248
/notes="putative"
2546..3349
/genes="Lhcb1*7"
2546..3349
/genes="Lhcb1*7"
/function="antenna complex"
/codon_start=1
/product="light harvesting chlorophyll a/b-binding
protein"
/protein_id="BAA25394.1"
/db_xref="GI:3036951"
/translation="MAAATMSLSPSFACKAVKLSYSSYEIINGKVTWRKTVIKAP
VGRNPMYGPDKVYKILGSGNSPSYLTGTFPGDYGWDTAGLSADPTFAKRELEVI
HGRNMLGALGVFPPELLARNGVKGEA*WFKAGSQIFSEGGLDVLGNLSLVHAQSL
ALWAQVVLGMGAVEGRVAGGIEGVDDPYFGGSDPLGLADDPAPFAELKVKKEIK
GLAMPSMFGFVQIAVTQKGFLENLADHLADPNNNAWAYATNFVPGK"
sig_peptide
2546..2647
/gene="Lhcb1*7"
/notes="putative"
mat_peptide
2648..3346
/gene="Lhcb1*7"
/product="light harvesting chlorophyll a/b-binding
protein"
/notes="putative"
ORIGIN
Query Match 10.0%; Score 100; DB 8; Length 3663;
Best Local Similarity 57.0%; Pred. No. 1,4e-08;
Matches 309; Conservative 0; Mismatches 210; Indels 23; Gaps 6;
QY 8 ACTTTGATTAATTAATTAATTAATTCGTATGACAAGGATTAACCTAGCTAATGATAATA 67
Db 1975 AATCCTGGTTCGAATTAATTAATTAATTTATGTAAGCAAGGGTTAATCTTAGCTAATAATG 1916
QY 68 ACTTCAGATCTTAATCAATTAATCAACAGCAATCACGGTCAATAGACAGCGCTTGAGAGA-AGATT 126

```

```

Db      1915 TTCCGAGTCTTAGTTTAAAGAACCAATTAAATACATATAGATTGGACGGACAGTAT 1856
QY      127  AATGTGATGTCATTCATATTTTCAAGATCATTAATGATAGGGA-----ATATCAAGCA 182
Db      1855 TGGTAACACATGATAAATGTTCTGGAAATCAGGAATAGTAATGTAGCATATAATATTC 1796
QY      183  ATAAATAACGATAAATGGCAATTAAGCTAAATGAAGGAGAAATGATTCACCCCAATATTGAATG 242
Db      1795 TTGAACACACATAAATGATATTTGATTAATAGGAGGAATGATCTTTTAATAAGGGA 1736
QY      243  AGGTGGATGATCTCTCTTTTGAACATGATGAATGATGGCAATAACTAGATGTTGGCA 302
Db      1735 GAATAAATGAACCTTCACCTGACCATGAT---TGATAGACAATCCTCGAATGATCGAG 1679
QY      303  CCCTTCTCGATCTAATGAAAAAGTATGGAATAGTATGATATCGAATCTCTTTAGAAAG 362
Db      1678 GATTTCTCGATCT-----GATGGTAAAGTGTAGACAAAATTTCTGTGAAAAG 1630
QY      363  GTAGTGAATGCTCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAAATATTTTATCAGAG 422
Db      1629 GTGATCTTCTGATCAAGTGAGAGCCGATTTGTTTGTAGTCAAGTGTCTTTTACAAA 1570
QY      423  AATATTACA---TCCGCTCTCTCCCTATCTCTTTTCTATTATATGAGGACAT-TCCTC 478
Db      1569 TGAATATCAGATGCCCTCATCTCTCTTTTATTTATTTATATGAGGACATGTTACT 1510
QY      479  AATCAATCTTAAAGTATACACACCAAGATATTCATATAAATATTTTGTGAATATCT 538
Db      1509 AGAAACCCCTAATATTACAAGTCAAGAAATATCCACTAGATATATCTTTAGTATCT 1450
QY      539  AT 540
Db      1449 AT 1448

```

RESULT 6

```

CR678065      164256 bp      DNA      linear      HTG 12-AUG-2004
LOCUS      Danio rerio clone DKEY-238D19, WORKING DRAFT SEQUENCE, 14 unordered
DEFINITION      pieces.

```

```

ACCESSION      CR678065
VERSION      CR678065.1 GI:51175972

```

```

KEYWORDS      HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE      Danio rerio (zebratfish)

```

ORGANISM

```

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 164256)

```

REFERENCE

```

AUTHORS      Burton, J.
TITLE      Direct Submission
JOURNAL      Submitted (11-AUG-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

```

COMMENT

```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK238D19
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 158159 bases at least Q40
Consensus quality: 159450 bases at least Q30
Consensus quality: 160402 bases at least Q20
Insert size: 162956; sum-of-contigs
Insert size: 176866; 2.6% error; agarose-fp
Quality coverage: 5.86x in Q20 bases; sum-of-contigs Quality
coverage: 5.66x in Q20 bases; agarose-fp
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 5677: contig of 5677 bp in length
* 5678 5777: gap of 100 bp
* 5778 18108: contig of 12331 bp in length
* 18109 18208: gap of 100 bp
* 18209 21518: contig of 3310 bp in length
* 21519 21618: gap of 100 bp
* 21619 64982: contig of 43364 bp in length
* 64983 65082: gap of 100 bp
* 65083 74563: contig of 9481 bp in length
* 74564 74663: gap of 100 bp
* 74664 78041: contig of 3378 bp in length
* 78042 78141: gap of 100 bp
* 78142 87675: contig of 9534 bp in length
* 87676 87775: gap of 100 bp
* 87776 111942: contig of 24167 bp in length
* 111943 112042: gap of 100 bp
* 112043 118229: contig of 6487 bp in length
* 118230 128112: contig of 9483 bp in length
* 128113 128212: gap of 100 bp
* 128213 135676: contig of 7464 bp in length
* 135677 135776: gap of 100 bp
* 135777 13672: contig of 2896 bp in length
* 13673 138772: gap of 100 bp
* 138773 151826: contig of 13054 bp in length
* 151827 151926: gap of 100 bp
* 151927 164256: contig of 12330 bp in length.

```

FEATURES

Source

```

1. 164256
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-238D19"
/clone_lib="DanioKey"
1. 5677
/note="assembly fragment:00119
fragment_chain:1"
5778..18108
/note="assembly fragment:00216
fragment_chain:1"
18209..21518
/note="assembly fragment:00097
fragment_chain:1"
21619..64982
/note="assembly fragment:01191
fragment_chain:1"
65083..74563
/note="assembly fragment:00372
fragment_chain:1"
74664..78041
/note="assembly fragment:00077
fragment_chain:1"
78142..87675
/note="assembly fragment:00293
fragment_chain:2"
87776..111942
/note="assembly fragment:00928
fragment_chain:2"
112043..118529
/note="assembly fragment:00164
fragment_chain:2"
118630..128112
/note="assembly fragment:00453
fragment_chain:3"
128213..135676

```

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature


```
QY 306 TTCTCGGATCTAATGAAAAAGATATCGAATAGTAGAATAATCGAATCTCTTTTGAAGGTA 365
|||
Db 6920 TTNTNTNNNNNAAATTAATAATTTTATANTANATAATATTAATNAATANNNNNAA 6979
|||
QY 366 GTGATGCTCTTTATCTAGAGAGAGAGCTGCTTTTCAAGATATTTTATCAGAGAT 425
|||
Db 6980 ATTAAAAANTATTATTTTTHAATATTTNNAAAAAATAATTTTATTTTATTTTAT 7039
|||
QY 426 ATTACATCCCTCTCTCCCTATCTCTTTTCTATTTATATGCGACATCTCCTCAATCAAT 485
|||
Db 7040 TTTTNTTAAATAAATAATATATATAATTTTATATAATNANNNTTNNANATTANT 7099
|||
QY 486 CCTAAAGTACATACACCAAGATATCTAATAAATATTTTGTGAATTTCTATATATA 545
|||
Db 7100 NTNTNTNTNTNTTNTTNTTAAAAAATAATTTTAAATAATTTTNTTANTTATANAAN 7159
|||
QY 546 AAATAGCTGTAGCACTCGCTCGTATTCGACTACTCGGTTCAGGACCTGTCA 605
|||
Db 7160 AAAATNTTATNTNTNAAAAANNNTNTTATATNTNTNTNTTTTNTTNTTNTT 7219
|||
QY 606 TTTTAACTAGCGCTCGATACATCACTTTCTACGATCTGCTTCATGCAAACTTAAT 665
|||
Db 7220 NNTNTTNNANANATNTNTNTNTNTTAAATTTTAAATAATAATTTATTTTAT 7279
|||
QY 666 GAAAGCAGATTTTGACCCATACATATATGACAAATTCCTCCAAAGAACATGGCT 725
|||
Db 7280 ATATATTTATATAAATAATATATATATAAATAATAAATAAATAAATAAATAA 7339
|||
QY 726 CTTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGAAATTTATTTAAGATAGT 785
|||
Db 7340 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7399
|||
QY 786 GTTTTCTTTCTTTCTTCTATCTAAGAGTAAGCAACCATGATAGAAAGCGCTTAG 845
|||
Db 7400 TATATATATATATATATATATATATATATATATATATATATATATATATATATA 7459
|||
QY 846 TAACTATATCAAGAAATGGTGTGTTTCTTAAATATGGAATAAAATTTGTGAATAT 905
|||
Db 7460 TATATATATAT--ATATATATATATATATATATATATATATATATATATACATAC 7516
|||
QY 906 AGAAGATTAGATCAAT 921
|||
Db 7517 ACACAATTAGTAAAT 7532
|||

RESULT 10
CR388040/c 156463 bp DNA linear HTG 09-APR-2004
LOCUS Danio rerio clone CH211-170D12, *** SEQUENCING IN PROGRESS ***, 11
DEFINITION unordered pieces.
ACCESSION CR388040
VERSION CR388040.1 GI:46357555
KEYWORDS HTG; HTGS PHASE1.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 156463)
            Sims,S.
            Direct Submission
            Submitted (09-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            ----- Project Information
            Center project name: zC170D12
            ----- Summary Statistics
```

Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 153457 bases at least Q40
Consensus quality: 153798 bases at least Q30
Consensus quality: 154184 bases at least Q20
Insert size: 155463; sum-of-contigs
Quality coverage: 7.69x in Q20 bases; sum-of-contigs Quality
coverage: 7.75x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 42348: contig of 42348 bp in length
* 42349 42448: gap of 100 bp
* 42449 58828: contig of 14380 bp in length
* 58829 59259: contig of 13031 bp in length
* 59260 70059: gap of 100 bp
* 70060 74154: contig of 4095 bp in length
* 74155 74254: gap of 100 bp
* 74255 83730: contig of 9476 bp in length
* 83731 83830: gap of 100 bp
* 83831 111539: contig of 27709 bp in length
* 111540 130084: contig of 18445 bp in length
* 130085 130184: gap of 100 bp
* 130185 138133: contig of 7949 bp in length
* 138134 138233: gap of 100 bp
* 138234 146836: contig of 8603 bp in length
* 146837 151721: contig of 4785 bp in length
* 151722 151821: gap of 100 bp
* 151822 156463: contig of 4642 bp in length.

FEATURES

source

1..156463
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-170D12"
/clone_lib="CHORI-211"
1..42348
/note="assembly fragment:01376"
fragment_chain:1
clone_end:SP6
vector_side:left
42449..58828
/note="assembly fragment:00607"
fragment_chain:1
58829..69959
/note="assembly fragment:00444"
fragment_chain:2
70060..74154
/note="assembly fragment:00013"
fragment_chain:2
74255..83730
/note="assembly fragment:00343"
fragment_chain:2
83831..111539
/note="assembly fragment:01016"
fragment_chain:2
111640..130084
/note="assembly fragment:00774"
fragment_chain:2
130185..138133
/note="assembly fragment:00248"
fragment_chain:2
138234..146836
misc_feature
1..42348
/note="assembly fragment:01376"
fragment_chain:1
clone_end:SP6
vector_side:left
42449..58828
/note="assembly fragment:00607"
fragment_chain:1
58829..69959
/note="assembly fragment:00444"
fragment_chain:2
70060..74154
/note="assembly fragment:00013"
fragment_chain:2
74255..83730
/note="assembly fragment:00343"
fragment_chain:2
83831..111539
/note="assembly fragment:01016"
fragment_chain:2
111640..130084
/note="assembly fragment:00774"
fragment_chain:2
130185..138133
/note="assembly fragment:00248"
fragment_chain:2
138234..146836
misc_feature


```

misc_feature 148165..160534
/notes="assembly_fragment:00424"
ORIGIN
Query Match 6.4%; Score 64; DB 2; Length 160534;
Best Local Similarity 34.7%; Pred. No. 0.021;
Matches 215; Conservative 0; Mismatches 401; Indels 3; Gaps 1;

QY 313 ATCTAATGAAAAGATGATGGAATAGTAGAATAATCGAATCTCTTTGAAAAGGTAGTGATG 372
D 3938 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 8997
QY 373 TCTTTATCTAGAGAGAAAGTCGCTCTTTCCTCAAGAAATATTTTATCAGAGAAATATACAT 432
D 8998 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9057
QY 433 CCCCTCTCTCCCTATCTCTCTTTCTATTTATATGCGACATTCCTCAATCAATCTCTAA 492
D 9058 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9117
QY 493 GTACATACACCAAGAAATATTCATATAAAATATTTTGTGAATATCTTATATATAAACTAG 552
D 9118 TTAATATTAATTAATTTTATTTTNNATTAATTTTAAATTAATTAATTAATTAATTAAT 9177
QY 553 CTGTT--AGCACTCGACCTCGCTCGCTGCTGCTACTCGGTTACGAGCCCTGCTCATTTA 609
D 9178 TTTTNTTAAATAATATAATANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9237
QY 610 CTAATCGACCTCGATTACATCACTCTCTACGAPACTGCTCAATGCTCAAAATCTTAATG 669
D 9238 TTTTNNNNNTTTNTTTTNTTNTNANATNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9297
QY 670 GCAGATTTGACCAATACATATATGACAAATTCGCTCCCAAGAAACATCGCTCTTA 729
D 9298 NANNNNNNNNNAATATAAATAAATAATATATATTTTTTTTTTACTTGAATTTTTTGT 9357
QY 730 TAGTGAATATCGTTAGCTGTTATAGAAGATCTGAATTTATTTTATAAGATAGTGTT 789
D 9358 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 9417
QY 790 TTTTCTTTCTTTTCATCTAATCAAGAGTAAAGCAACCATGAAPAGAAAAGGCTAGTAC 849
D 9418 TATATTGTTTNAITATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 9477
QY 850 TATATATCAAGAAATGGTCTTTTCTTTTAAATATGATGATAAAATTTGTGAATAGAA 909
D 9478 TTAATATTTATTTTATTTTATTTTATTTTATTTATATATATATATATATATATATAT 9537
QY 910 GATTAGATCAATTAACAAA 928
D 9538 ATATATATATATATATATA 9556

RESULT 12
AC067900
LOCUS Homo sapiens chromosome 4 clone RP11-775M3 map 4, WORKING DRAFT
DEFINITION SEQUENCE, 5 unordered pieces.
ACCESSION AC067900
VERSION AC067900.3 GI:12545341
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 201312)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Bouckghalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galligan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Leducque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,J.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Navlor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,E., Spencer,B.,
Stange-Thomann,N., Subramanian,A., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 201312)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Bouckghalter,B., Brown,A.,
Camparata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jehocky,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Navlor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 26, 2001 this sequence version replaced gi:8138718.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7524
Center clone name: 775 M.3
----- Summary Statistics
Sequencing vector: M13; M7815; 41% of reads
Sequencing vector: Plasmid; n/a; 59% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 198330 bases at least Q40
Consensus quality: 199387 bases at least Q30
Consensus quality: 200109 bases at least Q20
Insert size: 210000; agarose-gel
Insert size: 200912; sum-of-coverage
Quality coverage: 9.1 in Q20 bases; agarose-gel
Quality coverage: 9.5 in Q20 ba.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

	77971: gap of 100 bp	77971: contig of 77871 bp in length
*	77972	77971: gap of 100 bp
*	77972	83359: contig of 5188 bp in length
*	83160	83359: gap of 100 bp
*	83260	115460: contig of 32201 bp in length
*	115461	115560: gap of 100 bp
*	115561	158395: contig of 43435 bp in length
*	158096	159095: gap of 100 bp
*	159096	201312: contig of 42217 bp in length

```

FEATURES
* 159096 201312 Contigs of 4
  Location/Qualifiers
    1. 201312
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="4"
      /map="4"
      /clone="RP11-775M3"
      /clone_lib="RPC1-11 Human"
      /notes="assembly_fragment"
    1. 77871
      clone_end:SP6
      vector_side:left
      77972..83159
      /note="assembly_fragment"
      83260..115460
      /note="assembly_fragment"
      11561..158995
      /note="assembly_fragment"
      159096..201312
      /note="assembly_fragment"

```

ORIGIN

Query Match 6.4%; Score 63.6; DB 2; Length 201312;
Best Local Similarity 39.8%; Pred. No. 0.023;
Matches 371; Conservative 1; Mismatches 555; Indels 5;

[illegible]

Qy	481	TCATCCTTAAAGTACATACACCAGAATAATTCAATAAATAATTTTTTGGAATATTCTAT	540
Db	159074	NNNNNNNNNNNNNNNNNNNTTAATATATATAAATATATATAAATATATATAATAA	159133
Qy	541	TATAAAAACACTAGCTGTTAGCACTCGACCTCGTGGTATTTGACCTACTCGGTACGAGCCC	600
Db	159134	TATATAAATATATATAAATATATATAAATATATATAAATATATAAATATATAAATA	159193
Qy	601	TGTCATTTTACTAATCGACCTCGATTACATCACTTTCTACGATACTGCTTCATGTCAAATC	660
Db	159194	TATATAAATATATAAATATATATAAATATATAAATATATAAATATATAAATA	159253
Qy	661	TTAATGAAGCAGATTTTGACCCACATCAATATATGACAAATNGCTTCCAAGAAAAACA	720
Db	159254	TATATAAATATATATAAATATATAAATATATAAATATATAAATATATAAATATATA	159313
Qy	721	TGCGCTCT-TATAGTGAATAATCGTTAGACTGTTATAGAAGACTCGAATTTATTTTAAAG	779
Db	159314	TAAATATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATA	159373
Qy	780	AATAGTGTTTTTTCTTTTCTTTTCATATCTAAGGAGCTAAGCAACCATGAATAGAAAAG	839
Db	159374	TATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAA	159433
Qy	840	GCTTAGTAACATATAT--CAAAGGAATCGTGTTTTCTTTAAATATCGGATAAAAAATT	896
Db	159434	TATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAA	159493
Qy	897	TGTGAATATAGAAGATTAGATCAATTAACAAA	928
Db	159494	TATATATAAAGATGTGTTTATTTTAAAA	159525

RESULT 13

CR35440/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

SECRET

REFERENCE

[illegible]

JOURNAL.

PLANTAS

COMMENT

AUTHORS	Winden and van der J.
TITLE	Direct Submission
JOURNAL	Submitted (11-JAN-2001) Winden, van der J., Plant Molecular Genetics, Institute of Molecular Biology, Billrothstrasse 11, A-5020 Salzburg, AUSTRIA
FEATURES	Location/Qualifiers
source	1..4753
misc_feature	1..4753
ORIGIN	R
Query Match	6.3%; Score 62.8; DB 8; Length 4753;
Best Local Similarity	52.7%; Pred. No. 0.072;
Matches 277; Conservative	1; Mismatches 233; Indels 15; Gaps 6;
QY	5 TAGACCTTTTGATTAAATTAATTAATTTGATGACAAAGGATAAACCTAGTGAATGATA 64
DB	
DB	2298 TAAGCTTTATGTTAAACTAATTAATGAGNA--ATAGAGGGAGATCTCTAGTTAAACATA 2241
QY	65 ATAACTTCAGATCTATAATCAATTAACAGCAATACGGTCATAGCAGGTTGAGAGAAGA 124
DB	
DB	2240 ATAAATCCAGATAAAACCTGGGTATGTCAGCGATATAGACCGTGTTCATATGATGAT 2181
QY	125 TTAATGTGATGTCATTCGAATATTTTCAGATCATTAATGATAGGGGATATATCAAGCAAT 184
DB	
DB	2180 AAATATCGAATATGTGTTCAATGATTTGCAATGAATAAAGTAA---TAGAAGCATGCAAT 2124
QY	185 AAATAACGATAAATGGCAATTAAGTAAATAAGAGAGAAATGATTCACCC-AAATATTGAATGA 243
DB	
DB	2123 GAACTGAGAGATGACAATAAATGTAATAGAGAGAAATTTATCACCAAAATATTAGGTGA 2064
QY	244 GGTGATGATTTCTCTTTT--TGCAATGATGAATGATGGGCAAAATCTAGAAATGTTGGGA 302
DB	
DB	2063 CTGTGATGATTTCTCTTTCTCGTCGACAAACGCTGGATGATAGATAATGAAGATAGATAAG 2004
QY	303 CCCCTTCTCGGATCTAAATGAATAAGATAGTGAATAGTAGATAATCGAATCTCTTTAGAAG 362
DB	
DB	2003 GAATCTTTGGATCTGTGA-----ATGGCAGCTATATCGAAAATGTAATCAGTGAA 1951
QY	363 GTAGTGATGTTCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAAGAATATTTTATCAGAG 422
DB	
DB	1950 CAAGACACATTTTATATATTTCTATATTCAACCCITTCACATGCTCTTATCAAAA 1891
QY	423 AATATTACATCCCTCTCTCCCTATCTTTTCTTATTTATGCGACATTCCTCAATC 482
DB	
DB	1890 AGTGAAA--ATCCCTTTTGTCTTTATCTCTCTCTTATTTATAGGGGATATCCCTTGAA 1832
QY	483 AATCTTAAAGTACATACACAAAGATATTCAATAAAATATTTTTT 528
DB	
DB	1831 ACCCTAAAGTAGTCAACTGAAACAATATTCATGAATATTCCTT 1786
RESULT 15	AL592166
LOCUS	Human DNA sequence from clone RP11-269F19 on chromosome 1, complete sequence.
DEFINITION	AL592166
ACCESSION	AL592166
VERSION	AL592166.16 GI:29165241
KEYWORDS	HTG.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 90550)
TITLE	Wallis, J.
JOURNAL	Direct Submission
	Submitted (23-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 23, 2003 this sequence version replaced gi:18476714.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Swt, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/chr1>
RP11-269F19 is from the library RP11-11.1 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES	source	Location/Qualifiers	1. .90550	organism="Homo sapiens"	/mol_type="genomic DNA"	/db_xref="taxon:9606"	/chromosome="1"	/clone="RP11-269F19"	/clone_lib="RP11-11.1"	ORIGIN
Query Match			6.3%;	Score 62.6;	DB 9;	Length 90550;				
Best Local Similarity			43.1%;	Pred. No. 0.042;						
Matches 406;	Conservative	1;	Mismatches 525;	Indels	9;	Gaps	2;			
Qy	2	ATATAGACTTTTCATTAAATTAATTAATTTGATGACAAAGGATTAAACCTAGTTAATG	61							
Db	59484	ATATACATATATAAATAAATAAATAAAGTATTATGAATATATATATATAAATAT	5954							
Qy	62	ATAATAACTTCAGATCTATTAATCAATTAACAGCAATCACGGTCATGACGCGTTGAGAGA	121							
Db	59544	ATAAAAAATATATAAAATATATAAAATATATAAAATATAAAATATAAAATATAATAA	5960							
Qy	122	AGATTAAATGTGATGTCATTCGAATATTTCAAGATCATTAATCATGAGGGCAATATCAAGC	181							
Db	59604	AATATATATAAATATATAAATATATATAAATATATAAATATATAAATATATAAATA	5966							
Qy	182	AATAAAAATACGATAAATGGCATTTAAATAGTAATTAAGGAGAAATGATTCACCAATATTGAAT	241							
Db	59664	TATAAATATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAATATATA	5972							
Qy	242	GAGGTGGATGATCTTCTTTTTCGACATGATGAATGATGGCGAAATACTAGAAATGTTGGG	301							
Db	59724	AATATATAAATATATAAATAAATAAATATATAAATATATAAATATATAAATATATAA	5978							
Qy	302	ACCCCTTCTCGGATCTAATGAAAAAAGATGCGAATAGTAGAATATCGAATCTCTTTAGAAA	361							
Db	59784	TATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATA	5984							

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 06:24:54 ; Search time 493.658 Seconds
(without alignments)
10644.348 Million cell updates/sec

Title: US-10-009-966c-2_COPY_2155_3155

Perfect score: 1001

Sequence: 1 aatagagcttttgattaa.....accactactagaatccgg 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1000.2	99.9	4312	4 AAF25736	Aaf25736 Tobacco i
2	997	99.6	3294	4 AAF25735	Aaf25735 Tobacco i
3	997	99.6	4135	4 AAF25742	Aaf25742 Tobacco p
4	55.6	5.6	158001	12 ADL17884	Adl17884 Human pho
C 5	55.2	5.5	8056	8 ABZ10246	Abz10246 Haematopo
C 6	54.8	5.5	37515	6 ABO66998	Abog6698 Human ang
C 7	54.4	5.4	113515	6 ABL34175	Ab134175 Human imm
8	52.6	5.3	8201	6 ABL34511	Ab134511 Human met
9	52.6	5.3	8201	6 ABL70540	Ab170540 Chemicall
C 10	52	5.2	8056	8 ABZ10100	Abz10100 Haematopo
C 11	52	5.2	13584	6 ABL32615	Ab132615 Human imm
C 12	50.8	5.1	8056	8 ABZ10246	Abz10246 Haematopo
C 13	50.6	5.1	115218	8 ACA64845	Aca64845 Tumour su
C 14	50	5.0	5413	4 AAS46694	Aas46694 Tumour su
C 15	50	5.0	11745	6 ABK28332	Abk28332 DNA trans
16	49.4	4.9	11694	4 AAS46698	Aas46698 Tumour su
17	49.2	4.9	5360	6 ABL33937	Ab133937 Human imm
18	49.2	4.9	5360	6 ABL92297	Ab192297 Chemicall
19	49.2	4.9	5360	6 ABL43366	Ab143366 Human pol
C 20	49.2	4.9	5743	6 ABL34068	Ab134068 Human imm
C 21	49.2	4.9	7306	6 ABL33636	Ab133636 Human imm

C 22	49.2	4.9	34688	6 ABO67060	Abog67060 Human ang
23	49	4.9	9810	6 ABL32427	Ab132427 Human imm
24	48.8	4.9	6109	6 ABL32326	Ab132326 Human imm
25	48.8	4.9	6109	6 AAS61077	Aas61077 Human gen
C 26	48.6	4.9	50000	6 ABL56201	Ab156201 AnePV gen
27	48.4	4.8	12507	6 ABL32298	Ab132298 Human imm
28	48.4	4.8	20420	4 AAK73165	Aak73165 Human imm
C 29	48.4	4.8	20420	6 ABK66993	Abk66993 Human sec
30	48.4	4.8	37515	6 ABO66997	Abog6697 Human ang
31	48.2	4.8	7461	6 ABL33784	Ab133784 Human imm
32	47.8	4.8	15732	4 AAS45388	Aas45388 Chemicall
33	47.8	4.8	15732	6 ABK28333	Abk28333 DNA trans
34	47.8	4.8	115218	8 ACA64845	Aca64845 Haematopo
35	47.2	4.7	8056	8 ABZ10100	Abz10100 Haematopo
36	47.2	4.7	19965	4 AAK73166	Aak73166 Human imm
37	47.2	4.7	19965	6 ABK66932	Abk66932 Human sec
38	47	4.7	5572	6 ABL33424	Ab133424 Human imm
39	47	4.7	5572	6 ABL54359	Ab154359 Chemicall
40	47	4.7	5979	4 AAS45313	Aas45313 Chemicall
41	47	4.7	5979	6 ABK28152	Abk28152 DNA trans
42	47	4.7	11422	6 ABK39937	Abk39937 Human che
43	47	4.7	11422	6 ABL32219	Ab132219 Human imm
44	46.8	4.7	8047	6 ABN80324	Abn80324 Human che
45	46.8	4.7	14316	6 ABK31519	Abk31519 Signal tr

ALIGNMENTS

RESULT 1
AAF25736
ID AAF25736 standard; DNA; 4312 BP.
XX
AC AAF25736;
XX
DT 06-APR-2001 (first entry)
XX
DE Tobacco invertase promoter SEQ ID NO 2.
XX
KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.
XX
OS Nicotiana tabacum.
XX
PN NC0200077187-A2.
XX
PD 21-DEC-2000.
XX
PF 13-JUN-2000; 2000WO-DE001944.
XX
PR 12-JUN-1999; 99DE-02009998.
PR 04-APR-2000; 2000DE-02005992.
PR 26-APR-2000; 2000DE-02007494.
(ROIT/) ROITSCH T.
PI Roitsch T;
XX
DR WPI; 2001-080685/09.
PT New tapetum- and pollen-specific promoter from tobacco, useful for
preparing male sterile plants, particularly those with seedless fruits.
XX
PS Claim 6; Page 65-66; 74pp; German.
XX
CC This invention describes a novel nucleic acid (I) comprising a promoter
that is specific for tapetum and pollen. The invention also describes (a)
expression system containing at least one (I); (b) nucleic acid construct
(II) containing (I) plus at least part of an expressible nucleic acid
(III); (c) vector containing (I), the system of (a), or (II); (d) cells,
particularly plant cells, containing (II), the system of (a), (II) or the
vector of (c); (e) plants containing cells of (d); (f) seeds from the

Pat
Pending

CC cloning a promoter that is functionally homologous with (I). Constructs
 CC containing (I) and an invertase-encoding sequence are used to produce
 CC male-sterile plants (by co-suppression or antisense techniques) for
 CC preparation of hybrids, while constructs that contain (I) and a
 CC heterologous invertase-encoding sequence are used to prepare restorer
 CC plants (which allow propagation of the male-sterile plants). Seeds from
 CC these plants are used for in vitro embryogenesis of haploid or (double)
 CC diploid plants, and the new plants particularly produce seedless fruits.
 CC (I) can also be used to prepare transgenic plants that show increased or
 CC reduced production of endogenous materials, e.g. of plant hormones or
 CC proteins involved in provision of energy to developing tissue. (I)
 CC provide high level expression in a tissue- and time-specific manner, and
 CC do not require exogenous stimuli (e.g. temperature or chemicals) for
 CC regulation
 CC
 XX SQ Sequence 4135 BP; 1212 A; 774 C; 825 G; 1318 T; 0 U; 6 Other;

Query Match 99.6%; Score 997; DB 4; Length 4135;
 Best Local Similarity 99.6%; Pred. No. 8.6e-198;
 Matches 997; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AATATAGACTTTTGAATTAATAATTAATTTGATGACAAAGGATTAAACCTAGTTAAT 60
 DB |||||
 QY 61 GATATATACCTTCAGATCTATATCAATTAACAGCAATCAGGTCATAGCAGCGTTGAGAG 120
 DB |||||
 QY 1197 GATATATACCTTCAGATCTATATCAATTAACAGCAATCAGGTCATAGCAGCGTTGAGAG 1356
 DB |||||
 QY 121 AAGATTAAATGTCATGTCATTAATATTTCAAGATCATTAATGATAGGGGAATATCAAG 180
 DB |||||
 QY 1257 AAGATTAAATGTCATGTCATTAATATTTCAAGATCATTAATGATAGGGGAATATCAAG 1316
 DB |||||
 QY 181 CAATTAATAAGCATTAATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
 DB |||||
 QY 1317 CAATTAATAAGCATTAATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1376
 DB |||||
 QY 241 TGAGTGTGATGATCT 300
 DB |||||
 QY 1377 TGAGTGTGATGATCT 1436
 DB |||||
 QY 301 GACCTTCTCGATCTAATGAAAAAGTATGGAATAGTAGATGATGATGATGATGATGATGAT 360
 DB |||||
 QY 1437 GACCTTCTCGATCTAATGAAAAAGTATGGAATAGTAGATGATGATGATGATGATGATGAT 1496
 DB |||||
 QY 361 AGGTAGTGTGATGATCT 420
 DB |||||
 QY 1497 AGGTAGTGTGATGATCT 1556
 DB |||||
 QY 421 AGAATATTACATCCCT 480
 DB |||||
 QY 1557 AGAATATTACATCCCT 1616
 DB |||||
 QY 481 TCAATCTTAAGTACATGACCAAGATATTCATTAATTAATTAATTAATTAATTAATTAAT 540
 DB |||||
 QY 1617 TCAATCTTAAGTACATGACCAAGATATTCATTAATTAATTAATTAATTAATTAATTAAT 1676
 DB |||||
 QY 541 TATAAAATAGCTGTTAGCACTCGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 600
 DB |||||
 QY 1677 TATAAAATAGCTGTTAGCACTCGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1736
 DB |||||
 QY 601 TGTCAATTACTAATGACCTCGATTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
 DB |||||
 QY 1737 TGTCAATTACTAATGACCTCGATTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1796
 DB |||||
 QY 661 TTAATGAAAGCAGATTTTGACCCATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
 DB |||||
 QY 1797 TTAATGAAAGCAGATTTTGACCCATCAATTAATTAATTAATTAATTAATTAATTAATTA 1856
 DB |||||
 QY 721 TGGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGTAATTTATTTAAGA 780
 DB |||||
 QY 1857 TGGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGTAATTTATTTAAGA 1916
 DB |||||

QY 781 ATAGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 840
 DB |||||
 QY 1917 ATAGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1976
 DB |||||
 QY 841 CTTAGTAACATATATATCAAGGAATGGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCTTT 900
 DB |||||
 QY 1977 CTTAGTAACATATATATCAAGGAATGGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCT 2036
 DB |||||
 QY 901 AATATAGAGATTAGATCATCAATTAACAAAGGTTATCGTGGAGTGGTAAAGAGGGGACC 960
 DB |||||
 QY 2037 AATATAGAGATTAGATCATCAATTAACAAAGGTTATCGTGGAGTGGTAAAGAGGGGACC 2096
 DB |||||
 QY 961 TATGTCGTTATATAGAGGGGTACCCACTACTAGAAATCCGG 1001
 DB |||||
 QY 2097 TATGTCGTTATATAGAGGGGTACCCACTACTAGAAATCCGG 2137
 DB |||||

RESULT 4
 ADL17884
 ID ADL17884 standard; cDNA; 158001 BP.
 XX
 AC ADL17884;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human phosphotyrosyl phosphatase activator, PTPA, gene.
 XX
 KW Human; ds; antisense; phosphotyrosyl phosphatase activator; PTPA;
 KW hyperproliferative disorder; developmental disorder; infection;
 KW inflammation; tumour; gene.
 XX
 OS Homo sapiens.
 XX
 PN US2004023906-A1.
 XX
 PD 05-FEB-2004.
 XX
 PF 01-AUG-2002; 2002US-00211179.
 XX
 PR 01-AUG-2002; 2002US-00211179.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Dean NM, Dobie KW;
 XX
 DR WPI; 2004-132607/13.
 XX
 PS New antisense compound targeted to a nucleic acid molecule encoding
 PT phosphotyrosyl phosphatase activator, for modulating expression of
 PT phosphotyrosyl phosphatase activator or treating hyperproliferative
 PT disorders.
 XX
 PS Example 15; SEQ ID NO 11; 131bp; English.
 XX
 CC The invention relates to a compound 8-80 nucleobases in length targeted
 CC to a nucleic acid molecule encoding phosphotyrosyl phosphatase activator
 CC (PTPA), that specifically hybridises with the nucleic acid molecule
 CC encoding phosphotyrosyl phosphatase activator and inhibits the expression
 CC of phosphotyrosyl phosphatase activator, i.e. an antisense
 CC oligonucleotide. Also included are a composition comprising the compound
 CC and a pharmaceutical carrier or diluent, a method of inhibiting the
 CC expression of phosphotyrosyl phosphatase activator in cells or tissues, a
 CC method of treating an animal having a disease or condition associated
 CC with phosphotyrosyl phosphatase activator and a method of screening for
 CC an antisense compound. The disease or condition is a hyperproliferative
 CC disorder or developmental disorder. The compound, particularly the
 CC antisense oligonucleotide is useful in modulating the function of nucleic
 CC acid molecules encoding phosphotyrosyl phosphatase activator. The
 CC antisense compound can also be used as research tools and diagnostics. It
 CC can also be used as tools in differential and/or combinatorial analyses
 CC to elucidate expression patterns of a portion or the entire complement of
 CC genes expressed within cells and tissues. The compound can also be used
 CC for treating diseases or conditions associated with phosphotyrosyl

Query Match	5.6%	Score 55.6	DB 12	Length 158001
Best Local Similarity	42.6%	Pred. No. 0.085		
Matches 346	Conservative 1	Mismatches 460	Indels 5	Gaps 1
CC phosphatase activator, preferably hyperproliferative disorder or				
CC developmental disorder. The compound can also be used as prophylaxis,				
CC e.g. to prevent or delay infection, inflammation or tumour formation. The				
CC present sequence is the human PRPA gene which is a target for the				
CC antisense oligonucleotides of the invention.				
XX				
SQ Sequence 158001 BP; 34764 A; 40497 C; 41696 G; 40433 T; 0 U; 611 Other;				
117 AGAAGATTAAATGTGATCTGTCATTCACATATTTCAAGATCAATTAATGATAGGGAATAT 176				
22124 ATAA 22183				
177 CAAGCAATAATAACGATAAAGCCATTAAAGTAATAGGAGAAATGATTCACCAATAT 236				
22184 TAATATATAAT 22243				
237 TGAATGAGGTGGATGATCTCTCTTTTGACATGATGAATGGCAATACTAGATG 296				
22244 ATATATATAAAATAATATATAAATATATAAATATATATATATATATATATATAT 22303				
297 TTGGGAGCCCTTCGGATCTAATGAAAGATGATGAATAGTAGATAATCGAATCTCTTT 356				
22304 ATAAAT 22363				
357 AGAAGGTAGTGAATGTCTTTTATCTAGAGAGAACTGCTCTTTTCAAGAATATTTTA 416				
22364 ATAAATATATAAAATATATAAATATATATATATATATATATATATATATATATATA 22423				
417 TCAGAGAATATTACATCCGCCCTCTCCCTATCTCTTTTCTATTTATATGGGACATCC 476				
22424 TTATATATAAATATATAAATATATATATATATATATATATATATATATATATATA 22483				
477 TCAATCAATCCTAAAAAGTACATACACCAAGATTTTCAATAAAAAATTTTTTTGAATTT 536				
22484 TATATATTATATAAATATATAAAATATATATATATATATATATATATATATATATAA 22543				
537 CTATTTAABAACTAGCTGTAGCACTCGACCTCGGTCGYATTGACTACTCGGTTACGA 596				
22544 ATATATATATATATATATATAAATATATAAATATATAAATATATAAATATATAA 22598				
597 GCCTGTCTTACTAATCGACCTCGATTACATCACTTTCTACGATACCTGCTCATGTCA 656				
22599 ATATATAAAATATATAATATATATATATATATATATATATATATATATATATAA 22658				
657 AATCTTAATGAAAGCAGATTTGACCCATACAAATATGACAAATTCGCTTCCAAAGAA 716				
22659 AATATACAATATATATATATAAATATATAAATATATAAATATATAAATATATATAT 22718				
717 AACATGCGCTCTTATAGTGAATATCGTTAGACGTTTATAGAAAGATCTCGAATTTATTT 776				
22719 AATATATAATATATAAATAATATACAAATATATATATATATATATATATATATAT 22778				
777 AAGAATAGTGTTTTTTCTTTTCTTTTCTATCTAAGGAGTAAAGCAACCATGAATAGAA 836				
22779 ATAAAAATATATATATATAAATATATAAAAAATATACAATATATAAATATATAAATAT 22838				
837 AAGCCTTAGTAACCTATATATATCAAGGAATGGTGTTTTTTCTTTTAAATATGGATAAATTT 896				
22839 AATATATAATATATAAATAATATATATATATATATATATATATATATATATATATAT 22898				
897 TGTCAATATAGAAGATTAGATCAATTAACAAA 928				
22999 ATATAATATATAATATATAAATATATAAATATATAAATATATAAATATATACAT 22898				

RESULT 5

ABZ10246/c

ID ABZ10246 standard; DNA; 8056 BP.

XX

AC	ABZ10245;
XX	
XX	16-JAN-2003 (first entry)
DE	
XX	Haematopoietic cell proliferation disorder related DNA sequence #386.
XX	
KW	Human; haematopoietic cell proliferation disorder; cytostatic;
KW	gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW	cytosine methylation state; gene; ds.
XX	
XX	Homo sapiens.
OS	
XX	WO20027272-A2.
PN	
XX	
PD	03-OCT-2002.
XX	
PF	26-MAR-2002; 2002WO-EP003401.
XX	
PR	26-MAR-2001; 2001US-0278333P.
XX	
PA	(EPIC-) EPIGENOMICS AG.
XX	
PI	Berlin K, Braun A, Distler J, Gueutig D, Howe A, Mueller J;
PI	Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI	Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI	Schwöbe I, Ziebarth H;
XX	
DR	WPI; 2003-018942/01.
XX	
PT	Detecting and differentiating between hematopoietic cell proliferative
PT	disorders, comprises contacting a target nucleic acid with a reagent that
PT	distinguishes between methylated and non-methylated CpG dinucleotides.
XX	
XX	Claim 28; SEQ ID NO 386; 117pp; English.
PS	
XX	
CC	The present invention describes a method for detecting and
CC	differentiating between haematopoietic cell proliferative disorders
CC	associated with at least 1 gene and/or their regulatory regions in a
CC	subject. The method comprises contacting a target nucleic acid in a
CC	biological sample obtained from the subject with at least 1 reagent,
CC	which distinguishes between methylated and non-methylated CpG
CC	dinucleotides within the target nucleic acid. AB209861 to AB211118
CC	represent specifically claimed nucleotide sequences from the present
CC	invention. Oligonucleotides from the present invention can be used: for
CC	differentiating between healthy haematopoietic cells and proliferative
CC	disorder haematopoietic cells; for differentiating between acute
CC	lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC	determining the cytosine methylation state and/or single nucleotide
CC	polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC	related sequences and their complements; and as primers for the
CC	amplification of haematopoietic cell proliferation disorder related DNA
CC	sequences. The nucleotide sequences from the present invention can also
CC	be used for detecting a predisposition to, differentiation between
CC	subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC	haematopoietic cell proliferative disorders. The present method enables a
CC	highly specific classification of haematopoietic cell proliferative
CC	disorders allowing for improved and informed treatment of patients
XX	
SO	Sequence 8056 BP; 37111 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

	Query Match	5.5%	Score 55.2	DB 8	Length 8056
	Best Local Similarity	42.8%	Pred. No. 0.067		
	Matches 395	Conservative 1	Mismatches 514	Indels 12	Gaps 2
Qy	3	TATAGACTTTTGATTAAATTAATTAATTTGTTATGACAAAGGATTAACCTAGTTAATGA	62		
Db	2610	TATAAATTTTTTTTTTCATTTTATTTTTATTTTTTAAAAATTTTTATTTAAATTTTAAATTT	2551		
Qy	63	TAATAACCTTCAGATCATATATCAATTAACAGCAATCACGGTCATAGCAGCGTTGAGAGAA	122		
Db	2550	TTTTAAATTTATTTTAAATTTAATTTATTTTTTTTTTTTTCATTAATTTTAAATTTAT	2491		
Qy	123	GATTAAATGTGATGTCYCATTTCAATCAATTTTCAAGATCATTAATGATGAGGGGAATATCAAGCA	182		

RESULT 5
ABZ10246/C
ID ABZ10246 standard; DNA; 8056 BP.
XX

QY	633	TTTCTACGATACTGCTTCATGTCGAATCTTAAATGAAAGCAGAGATTTTGACCCATACATAA	692			
DB	3867	AAACAAATAATACAAATTAATAATAAAACATTTAAATAAATAACATATACATACTAACTTA	3808			
QY	693	TATGACAAAATTGGTTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGTTAGACTGTT	752			
DB	3807	ACAAAACCTTCTAAATTAAATAACAACAATAATATATATATTACTTTAATTAAACATCC	3748			
QY	753	ATAGAAAGATCTGAATTTATTATAGAAATAGTGTGTTTTTTCTTTTCTTTTCATATCTAA	812			
DB	3747	AAAAATATACCTATTTTTTATAAATAATATTTATTTATATAAAAATATTTAAACAATTCAC	3688			
QY	813	GGACTAAAGCAACCATGAATAGAAAA	838			
DB	3687	AAACTCTCAAAACCATTAATTATATA	3662			
RESULT 8						
ABL34511						
ID	ABL34511	standard; DNA; 8201 BP.				
AC	ABL34511;					
DT	26-MAR-2002	(first entry)				
XX	Human metastasis associated gene SEQ ID NO: 64.					
DE	Metastasis associated gene; cytostatic; gene therapy; cancer;					
KW	cytosine methylation; gene; ds.					
XX	Homo sapiens.					
OS	WO200177376-A2.					
XX	18-OCT-2001.					
DD	06-APR-2001; 2001WO-EP003970.					
XX	06-APR-2000; 2000DE-01019058.					
PR	07-APR-2000; 2000DE-01019173.					
PR	30-JUN-2000; 2000DE-01032529.					
PR	01-SEP-2000; 2000DE-01043826.					
XX	(EPIG-) EPIGENOMICS AG.					
PA	Olek A, Piepenbrock C, Berlin K;					
XX	WPI; 2002-010922/01.					
XX	New nucleic acid derived from chemically treated metastasis genes, useful					
PT	for diagnosis of cancers by analysis of cytosine methylation, also for					
PT	treatment.					
XX	Claim 1; SEQ ID NO 64; 23pp + Sequence Listing; English.					
PS	The present invention provides a number of human metastasis associated					
CC	genes which are modified by cytosine methylation. The sequences can be					
CC	used in the diagnosis and treatment of cancer. The present sequence is					
CC	one of the genes of the invention					
XX	Sequence 8201 BP; 1609 A; 455 C; 2470 G; 3667 T; 0 U; 0 Other;					
SQ	Query Match 5.3%; Score 52.6; DB 6; Length 8201;					
Best Local Similarity 53.1%; Pred. No. 0.23;						
Matches 112; Conservative 0; Mismatches 99; Indels 0; Gaps 0;						
QY	689	ATAATATGACAAATTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGTTAGAC	748			
DB	4834	ATAAAATAAAAATAATAATAAAAATAACGTAATAGTTAGTGAATATTTATTT	4893			
QY	749	TGTTATAGAAGATCTCAATTTATTTATAGAATAGTGTGTTTTTTCTTTTCATAT	808			
DB	4894	TGTTATAGATTTGAGCAATGTTTTATGTTGATTTTTTTTTTCGTAATTTAATTATGT	4953			

QY 809 CTAAGGAGTAAGCAACCATGATAGAAAAGGCTTAGTAATATATATATCAAGGAATGTT 868
 DB 4954 TGGGATTGTAAGTTATATAGTATATATTTGTTGTAAGTTATTTTATTTTAAAGGAGTAAT 5013
 QY 869 GTTTTTCTTTAAATATGATGATAAAAATTTGT 899
 DB 5014 TTAGTTTTTTTATTTTGAATTTAGAAATGTT 5044

RESULT 9
 ABL70540
 ID ABL70540 standard; DNA; 8201 BP.
 AC ABL70540;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Chemically treated cell signalling DNA sequence complementary to#215.
 XX
 KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
 KW tumour; cytostatic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200202807-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-EP007471.
 XX
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-154758/20.
 XX
 PT Nucleic acid, useful for diagnosis and therapy of diseases associated
 PT with cell signaling e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with cell signaling.
 XX
 PS Claim 1; SEQ ID NO 430; 24pp + Sequence Listing; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of at least
 CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling. Note: The
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office
 XX
 SQ Sequence 8201 BP; 1609 A; 455 C; 2470 G; 3667 T; 0 U; 0 Other;
 Query Match 5.3%; Score 52.6; DB 6; Length 8201;
 Best Local Similarity 53.1%; Pred. No. 0.23;
 Matches 112; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 QY 699 ATAATATGACAAATGCTTCCAAGAAACATGCGCTTATATAGTAATCGTTAGAC 748
 DB 4834 ATAAAATAAAAAATATAATAAAAAAATACGTTAATAGTTTGTGTAATATTTATT 4893

QY 749 TGTATAGAAAGATCTGAATTTATTTATAGATAGTGTCTTTTCTTTCTTTTCATAT 808
 DB 4894 TGTATAGATATTAGGAATGTTTATATCTTGATATTTTTCGTAATTAATATGT 4953
 QY 809 CTAAGGAGTAAGCAACCATGATAGAAAAGGCTTAGTAATATATATCAAGGAATGTT 868
 DB 4954 TGGGATTGTAAGTTATATAGTATATTTGTTGTAAGTTATTTTATTTTAAAGGAGTAAT 5013
 QY 869 GTTTTTCTTTAAATATGATGATAAAAATTTGT 899
 DB 5014 TTAGTTTTTTTATTTTGAATTTAGAAATGTT 5044

RESULT 10
 ABZ10100/C
 ID ABZ10100 standard; DNA; 8056 BP.
 XX
 AC ABZ10100;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #240.
 XX
 KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW Gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277272-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-EP003401.
 XX
 PR 26-MAR-2001; 2001US-0278333P.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwöpe I, Ziebarth H;
 XX WPI; 2003-018942/01.
 XX
 PT Detecting and differentiating between haematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 XX
 PS Claim 28; SEQ ID NO 240; 117pp; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative

[illegible]

RESULT 11	
ABL32615	
ID	ABL32615 standard; DNA; 13584 BP.
XX	
AC	ABL32615;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Human immune system associated gene SEQ ID NO: 588.
XX	
KW	Human; immune system disease; cytosine methylation; antiasthmatic;
KW	antiarteriosclerotic; antianaemic; cytosstatic; nootropic;
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW	ds.
OS	Homo sapiens.
XX	
PN	WO200200928-A2.
XX	
PD	03-JAN-2002.
XX	
PF	02-JUL-2001; 2001WO-BP007537.
XX	
PR	30-JUN-2000; 2000DE-01032529.
PR	01-SEP-2000; 2000DE-01043826.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI; 2002-130909/17.
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful for
PT	diagnosis and treatment of diseases associated with abnormal cytosine
PT	methylation.
XX	
PS	Claim 1; SEQ ID NO 588; 32pp + Sequence Listing; German.
XX	
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention
XX	
SQ	Sequence 13584 BP; 3905 A; 112 C; 2520 G; 7047 T; 0 U; 0 Other;
Query Match	5.2%; Score 52; DB 6; Length 13584;
Best Local Similarity	46.6%; Pred. No. 0.34;
Matches	264; Conservative 1; Mismatches 296; Indels 5; Gaps 3;
Qy	1 AATATAGACTTTTGATTAATAATTAATTAATGTTATGACAAAGGATTAACCTAGTTAAT 60
Db	12840 ATTAGTGAATATGTAATTTTGTGTAATTTGAACAGAAAATGGGAATATT 12899
Qy	61 GATAATAACTCAGATCTATAATCAATTAACAGCAATCACGGTCATACACGGTTGAGAG 120
Db	12900 TTGTAGATATAAGATTTTATTATTAGTTATAGTAAATTATGTTAGTTATATCTTTGTGTT 12959
Qy	121 AAGATTAAATGTGATGTCATTCAT--ATTTCAGATCAATTATGATACGGGAATATC 177
Db	12960 GTGGAAATAAATATATTATTGAATGGAATAGAAAGTTGAGAAATATGTTTGAATTT 13019
Qy	178 AAGCAATAAATACGATAAATGGCAATTAAGTAAATAAGGAGAATGATTCACCCAAATTT 237
Db	13020 GATATATATAGAGTTACTGGTATTAGAAAAGATGGATTATTATTAATTAATGGTGT 13079

Qy	238	GAATGAGGTGGATGATCTCTCTTTTGCACAATGATGCGGCAAACTACTAGAACTGT	297
Db	13080	GAGATAATTGGTTATTATTGTAGAAATAAAATTAGATTTTATTATTTATAGTAAATTA	13139
Qy	298	TGGGACCCCTCTCCGATCTAAATCAAAAAAGTAGGAATAGTAG-ATAATCGAATCTCTTT	356
Db	13140	TATATATAATTTTTTAAATGATTAAAGGTTTAAATAGAAAAAGTAAAAATTTTAAATTTTT	13199
Qy	357	AGAAAGGTAGTGATTCCTCTTTTATCTAGACAGAAAGCTGCTTTTCAAAGATATTTTATA	416
Db	13200	AGAAGAAAAAGTATGTTGTATAATTTTGGAAATAAAGATATTTTTTAAATAGTTATATA	13259
Qy	417	TCAGAGAAATATTACATCCCCCTCTCCCTATCTCTCTTTCTTATTATA-TGGGACATTC	475
Db	13260	AATGCTGATATTATAAAGGAAACGATTGATAAATTTATTTTATTAAAAATAAAAATGCG	13319
Qy	476	CTCAATCAATCTTAAAGTACATACACCAAGATATTCGAATAAATATTTTTTTGAAAT	535
Db	13320	AGATATTATAATAAAGTGAAGATGAATTTAGATTGAATGAAGTATTTGTTAGTTAA	13379
Qy	536	TCTATTATAAAAACTAGCTGTTAGCA	561
Db	13380	GTTTTAGAAAAAGATTAAATATTAGAA	13405

RESULT 12
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.
XX
XX
AC
AC
XX
XX
DT (first entry)
DE
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
XX Human: haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
XX
OS Homo sapiens.
XX
XX W0200277272-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 26-MAR-2002; 2002WO-EP003401.
PF
XX
XX 26-MAR-2001; 2001US-0278333P.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorian P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet
PI Schwöpe I, Ziebarth H;
XX
XX MPI; 2003-018942/01.
DR
XX
XX Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent th
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
XX Claim 28; SEQ ID NO 386; 117pp; English.
PS
XX
XX The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for

CC	differentiating between healthy haematopoietic cells and proliferative
CC	disorder haematopoietic cells; for differentiating between acute
CC	lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC	determining the cytosine methylation state and/or single nucleotide
CC	polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC	related sequences and their complements; and as primers for the
CC	amplification of haematopoietic cell proliferation disorder related DNA
CC	sequences. The nucleotide sequences from the present invention can also
CC	be used for detecting a predisposition to, differentiation between
CC	subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC	haematopoietic cell proliferative disorders. The present method enables a
CC	highly specific classification of haematopoietic cell proliferative
CC	disorders allowing for improved and informed treatment of patients
XX	
SQ	Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
	Query Match 5.1%; Score 501.8; DB 8; Length 8056;
	Best Local Similarity 44.1%; Pred. No. 0.55;
	Matches 410; Conservative 2; Mismatches 494; Indels 24; Gaps
QY	11 TTTGATTAAATTAATTAATTTCTATGACAAAGGATTAAACCTAGTTAATGATAATAACT 70
DB	1359 TTTTATTATAAAATTAATTTTATATAAAATAAATTCGATTTTTTTTATTTTTT 1418
QY	71 TCAGATCTATAATCAATTAACAGCAATCAGGTCATAGCGGTTGAGAGAAGATTAAAT 130
DB	1419 TTAATTTTTTTTAAAGAAAAATAAAAAATTAATTTTTTAATAAAATTAATTAATA 1478
QY	131 GTGATGTCATTCGAATATTTCAAGATCAATTAATGATAGGGATATCAAGCATAATAA 190
DB	1479 TAATTAATTAATTAATTAATTAATTAATAATAATAAATAATAAATGTTAAAAAATTT 1538
QY	191 CGATAAATGCATTAAGTAATAATGAGAGAAATGATTCACCCAAATTTGAATGAGGTGGAT 250
DB	1539 TAATAAAATTAATTTTTTAAAGAAATTAATTTAAATTTAAATTTTATATTTAT 1598
QY	251 GATCTCTTTTTTGACAATGATGAATGA-----TGGCAAAATCTAGAACTTCGGAC 303
DB	1599 AATTTTAATATTTTAAAAAATTCGAATAAATGAATTTGTAAAATTTAAAAAATAATTTTA 1658
QY	304 CCTCTCGGATCTAANTGAAAGAAAGTAGTGAATAGTAGATAATCGAATCTCTTTAGAAAGG 363
DB	1659 ATTAATATGATAAAATTTTATTTTTTAATAATTAATAAAATTAATTAATAATAATA 1718
QY	364 TAGTGATGTGCTTTTATCTAGAGAGAAGTCCTCTTTTCAAAGAAATTTTTTATCAGAGA 423
DB	1719 TATATTTATAATTTTAATTTATTTAAAAA--ATTAAAAATAAAAAAATTTTTTAAATTAATA 1776
QY	424 ATATTACATCCCCTCTCTCCCTATCTCTTTTCTATTTATATATGGGACATTCCTCAATCA 483
DB	1777 AATTAAAAAATTAATATGTATTATATTTTAAATTAATAATAAAATTAATTTGTTTTA 1836
QY	484 ATCCTAAAAAGTACATACACCAAGAATATTCAATAAAAAATTTTTTTTGAATATTCTATTAT 543
DB	1837 AAAAAAATAAAAAAATAAAAAAATAATTTAAATTAATTTATTTTTTTTATTTAT 1896
QY	544 AAAAACTAGCTGTAGACATCTGACCTCGGTCGCTATGTACTCTCGGTACGAGCCCTGT 603
DB	1897 AAAAAATAAAATTTTTTAAAAAATAAAAAATAATTTAAATTAATAATAAAAAAATAAT 1956
QY	604 CATTACTAATCCAGCTCGAATTACATCACTTTCTAGATACCTGCTTCATGCTCAAACTCTTA 663
DB	1957 AATTTTTTAAAAATAAAATTAATTAATTAATTTTA-----TTTAAATTTTAATAAAA 2008
QY	664 ATGAAGACAGATTTTGACCCCATTAATAATAATGACAAAAATTTGCTTCCAAAGAAAAATGG 723
DB	2009 TTTTAAATAATTTTAAAAATAAAAAATAATTTTTTAAATATTTTTTTTAAAAAATTTTA 2068
QY	724 CTCCTTAGTGAATATCGTTTACACTGTTTATAGAAAGATCTGAAATTTATTTATAAGAAATA 783
DB	2069 TT-----AAAAAATATTTAAATATTAATAATAATAATTAATAATTTTATAATAATT 2121
QY	784 GTGTTTCTTTCTTTTCATCTATCAAGGATGAAGCAACCATGAATGAAAAAGCCTT 843

Qy	429	ACATCCCCCTCTCTCCCTATCTCTTTTCTATTATATGGGACATCTCTCAATCAATCCT	488
Db	17357	ATATTATTATTATTATTATTATTATTATTATTATAAATATATATATGATATATTAT	17298
Qy	489	AAAAGTACATACACCAAGAATATTCAAATAAATAATTTTTTGAATATCTATTATAAAAA	548
Db	17297	ATTATATGATA----TATTATATGATATATTATATTATATATATATATATAT	17242
Qy	549	CTAGCTGTGTAGCACTCGACCTCGGTGCTGATTTGACTCTCGGTTACGAGCCCTGTCAATT	608

609 ACTAATCGACCTCGATTACATCACTTTCACGATCTGCTTCATGTCGAATCTTAATGAA 668

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible][illegible]

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

T O U R I S M

D C

QY 203 AGATTAGATCAATTA 223

UD 16881 AATAATAATAATA 1686 /

RESULT 14

AAS46694
TD AAS46694 standard: DNA: 5413 BP

XX	
XX	
AC	
XX	
DT	
AAS46694;	
18-DEC-2001	
(5:rst entrv)	

XX
DE Tumour suppressor gene derived chemically modified semience #417

XX	
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
KW	tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW	cytosine methylation; ds.
XX	
OS	Homo sapiens.
XX	
FN	WO200169912-A2.
XX	
PD	20-SEP-2001.
XX	
PF	15-MAR-2001; 2001WO-EF002955.
XX	
PR	15-MAR-2000; 2000DE-01013847.
PR	06-APR-2000; 2000DE-01019058.
PR	07-APR-2000; 2000DE-01019173.
PR	

24 30-JUN-2000; 2000DE-01032325;
PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPIG-) EPIGENOMICS AG.

AA
PI Olek A, Piepenbrock C, Berlin K;
vv

DR WPI; 2001-602752/68.
yy

PT Fragments of chemically modified genes associated with tumor suppressor

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 07:44:14 ; Search time 3107.21 Seconds
(without alignments)
11739.223 Million cell updates/sec

Title: US-10-009-966C-2_COPY_2155_3155
Perfect score: 1001
Sequence: 1 aaatagacttttgattaaa.....accactactagaatccgg 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_btc.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	66.2	6.6	1101	9	CNS0039G
c 2	66	6.6	1101	9	CNS0039G
c 3	63.2	6.3	1101	9	CNS0039G
c 4	62	6.2	1101	9	CNS0039G
c 5	60.2	6.0	1092	9	CNS020K7
c 6	59.4	5.9	1101	9	CNS0039G
c 7	59	5.9	1542	9	AG386981
c 8	58.8	5.9	996	9	CNS0039G
c 9	58.8	5.9	2015	9	AG311015
c 10	58.6	5.9	928	9	CNS0039G
c 11	58.4	5.8	1132	8	BZ695089
c 12	58.4	5.8	1335	9	CT646802
c 13	58	5.8	1101	9	CNS0039G
c 14	57.4	5.7	1025	9	CNS014J2
c 15	57.4	5.7	1201	9	CNS01522
c 16	57	5.7	1225	9	CNS0161D
c 17	56.8	5.7	1067	9	AG525799
c 18	56.6	5.7	774	8	BZ517772
c 19	56.4	5.6	815	9	AG056601
c 20	56.4	5.6	1202	8	CC264481
c 21	56.2	5.6	1092	9	CNS020K7
c 22	56	5.6	442	8	BZ518729
c 23	56	5.6	1200	9	CNS016CO
c 24	56	5.6	1758	9	CU509408

25	55.2	5.5	938	9	CNS006TU	AL065906 Drosophil
26	55	5.5	959	9	CNS004RY	AL055406 Drosophil
c 27	55	5.5	945	9	CNS04DOK	AL285149 Tetraodon
c 28	55	5.5	1101	9	CNS00FEV	AL071298 Drosophil
c 29	54.8	5.5	1227	9	AG430010	AG430010 Mus muscu
c 30	54.6	5.5	815	9	AG471644	AG471644 Mus muscu
c 31	54.4	5.4	681	8	BZ051088	BZ051088 jnr62c01
c 32	54.4	5.4	786	8	BH58486	BH58486 BOMNU18TF
c 33	54.4	5.4	1094	9	CNS012FZ	AL101513 Drosophil
c 34	54.4	5.4	1101	9	CNS0087Y	AL069057 Drosophil
c 35	54.4	5.4	1101	9	CNS0100X	AL098379 Drosophil
c 36	54.2	5.4	1101	9	CNS0106X	AL098595 Drosophil
c 37	54.2	5.4	1227	9	AG430010	AG430010 Mus muscu
c 38	54.2	5.4	1268	9	AG347098	AG347098 Mus muscu
c 39	54	5.4	1167	9	CNS07360	AL427102 clone BA0
c 40	54	5.4	1204	9	CNS016E2	AL106628 Drosophil
c 41	54	5.4	1539	9	AG340947	AG340947 Mus muscu
c 42	54	5.4	2004	9	AG387068	AG387068 Mus muscu
c 43	53.8	5.4	590	5	BQ475857	BQ475857 platystom
c 44	53.8	5.4	1101	9	CNS00807	AL069440 Drosophil
c 45	53.6	5.4	210	5	BQ242105	BQ242105 Tae15034B

ALIGNMENTS

RESULT 1
LOCUS CNS0039G/c
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BAC08X10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063921
VERSION AL063921.1 GI:4941778
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 6.6%; Score 66.2; DB 9; Length 1101;
Best Local Similarity 16.8%; Pred. No. 0.0027;


```

Db 40 ATATATAATTTANATATNTANANTATAAANTNTNTANANTNTNTANANTNTNTATATAAANAANTCTN 99
Qy 62 ATAATAAAGTCCAGATCTATATCAATTAACAGCAATCACGGTCTACAGCGTTGAGAGA 121
Db 100 TCTATATAGATATATATATTTNTNTNTNTATATATANTNTANANANNTTATNTATATA 159
Qy 122 AGATTAATGATGATGTCATTCATCAATATTCAGATCATTATGATAGGGAATATCAAGC 181
Db 160 TNANTATANTANANANANANANATATATANTNTNTNTNTNTNTNTATAGKGNWNWATA 219
Qy 182 AATAAATAACGATAAATGGCAATTAAGATAAATAAGGAGAAATGATTCACCCCAATATTGAAT 241
Db 220 TANATATANTNTNTNTNTNTNTNTATATATATANTNANANANATATANANANTNAPATAN 279
Qy 242 GAGTGGATGATCTCTCTTTTGCAATGATGAATGATGGCAATATCTAGAAATCTGGG 301
Db 280 TANANATAGWGWATTTATTTATATAAATAAATAAATAATATATATATATATAAATTATA 339
Qy 302 ACCCTTCGCGATCTAATGAAATAAGTAT-GGAATAGTAGATAATCGAATCTCTTTAGAA 360
Db 340 TAATTAATTAATTTAAAWAAWAAATTTWATTTATATNAAAAATTTTATATATATAAAA 399
Qy 361 AGGTAGTGAATGCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAATATTTTATATCAG 420
Db 400 TAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 459
Qy 421 AGAATATATACATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 480
Db 460 ATATATTTATTTTAWAAWATTTTATATATATATATATATATATATATATATATATATA 519
Qy 481 TCAATCTTAAAGTACATACACCAAGAAATATTCATATAAATAATTTTTCGAATATCTTAT 540
Db 520 TAAATTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 579
Qy 541 TATAAATACTAGCTGTTA 558
Db 580 TATWAAAAAATAAATAA 597

-RESULT 7
AG386981/c
LOCUS 1542 bp DNA linear GSS 03-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-201G10.TJ, genomic survey
sequence.
ACCESSION AG386981
VERSION AG386981.1 GI:47998186
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
2 (bases 1 to 1542)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsr.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Clones are derived from the mouse BAC library MSMg01. For BAC
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY

```

```

Vector : pBACE3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
Location/Qualifiers
source 1..1542
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-201G10.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 5.9%; Score 59; DB 9; Length 1542;
Best Local Similarity 43.8%; Pred. No. 0.079;
Matches 420; Conservative 0; Mismatches 525; Indels 15; Gaps 4;

Qy 6 AGACTTTTCATTAAATTTAATTTAATTTGATGACAAAGGATTAAACCTTAGTTAAATGATAA 65
Db 1075 AAAAAATATATATATATATATATATATATATATATATATATATATATATATATATATA 1016
Qy 66 TAACTTCAGATCTATATCAATTAACAGCAATCACGGTCTACAGCGTTGAGAGAAAT 125
Db 1015 ATAAAAAATAAATATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 956
Qy 126 TAAATGATGATGTCATTCATATATTTCAAGATCATTAAATGATAGGGGAATATCAAGCAATA 185
Db 955 ATAATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 896
Qy 186 AATAACGATAAATGCGATTAAAGTAAATTAAGGAGAAATGATTCACCCCAATATTGAATGAGG 245
Db 895 TAAATTTAATTAATATATATATATATATATATATATATATATATATATATATATATATAT 836
Qy 246 TGGATGATCTCTCTTTTTCAGCAATGATGAATGATGGGCAATCTAGATGATGTTGGGACCC 305
Db 835 AATATTATAAATAATTAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 776
Qy 306 TTCTCGGATCTAATGAATAAAGTATGGAATAGTAGATATCGAATCTCTTTTGAAGAAGGTA 365
Db 775 TAAAAATATNAAAAATNAAAAATNAAAAATNAAAAATNAAAAATNAAAAATNAAAAATNAAA 716
Qy 366 GTGATTGCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAATAATTTTATATCAGAAAT 425
Db 715 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 656
Qy 426 ATTAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAATCAAT 485
Db 655 AATAAATAAATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 596
Qy 486 CCTAAAGTACATACACCAAGAAATATTCATATAAATAATTTTTCGAATATTTCTATATAA 545
Db 595 ATTAAATNTAATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 542
Qy 546 AAATAGCTGTTAGCACTCGACCTCGTGGTATTTGACTACTCTCGGTTCAGGACCCCTGTCA 605
Db 541 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 482
Qy 606 TTTTACTAATCGACCTCGATTACATCACTTTTACGATCTGCTTCTCATGCAATCTCTAAT 665
Db 481 TTTTATTTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 422
Qy 666 GAAAGCAGATTTTGACCCCATACAAATAATGACAAAAATTCCTCCAAAGAAACATGGCT 725
Db 421 AAA---ATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 365
Qy 726 CTTATAGTCAATATCGTTAGACTGTTATAGAAAGATCTGAATTTTATATATAGATAGT 785
Db 364 ATATTATTAAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 305
Qy 786 GTTTTTTTCTCTTTT---CTTTTCATATCTAAGGAGTAAAGCAACCATGAATGAAAAAGCC 841

```

```

Db      304 TTATTTTAAATTAATAATATATATTAATTTAAATAATTAATAATTAATTTAAATAATAATA 245
QY      842 T--TAGTAATACTATATCAAAAGGAATGGTGTGTTTTTCTTTAAATATGATAAAATTTGT 899
Db      244 TAAATTTAAATTAATTTAAATTAATTAATTAATTTTAAATAATATATATTAATAA 185
QY      900 GAATATAGAAGATTAGATCAATTAACAAAGGTTATGGTGGAGTGGTGAAGCAGAGCGGAC 959
Db      184 ATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125

RESULT 8
CNS00FUH      996 bp      DNA      linear      GSS 03-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence TET3 end of BAC:
DEFINITION      BACR31021 of RPCL-98 library from Drosophila melanogaster (fruit
SOURCE      fly), genomic survey sequence.
VERSION      AL071063.1 GI:49511105
KEYWORDS      GSS.
ORGANISM      Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
REFERENCE      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
AUTHORS      Ephydroidea; Drosophilidae; Drosophila.
TITLE      1 (bases 1 to 996)
JOURNAL      Genoscope.
COMMENT      Direct Submission
      Determination of this BAC-end sequence was carried out as part of a
      collaboration with the Berkeley Drosophila Genome Project (BDGP).
      The BDGP is constructing a physical map of the Drosophila
      melanogaster genome using these BACs. For further information
      please see http://www.fruitfly.org The BDGP Drosophila
      melanogaster BAC library was prepared by Kazutoyo Osoegawa and
      Aaron Mammosier in Pieter de Jong's laboratory in the Department of
      Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
      NY. The library is named RPCL-98 and was constructed by partial
      EcoRI digestion of Drosophila DNA provided by the BDGP from the
      isogenic strain y2; cn bw sp, the same strain used for the BDGP's
      P1 and EST libraries. A more detailed description of the library
      and how to order individual BAC clones, the entire library, or
      filters for hybridization from the BACPAC Resource Center can be
      found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
      Location/Qualifiers
      1..996
      /organism="Drosophila melanogaster"
      /mol_type="genomic DNA"
      /db_xref="taxon:7227"
      /clone="BACR31021"
      /clone_lib="RPCL-98"
      /note="end : TET3"

ORIGIN
Query Match      5.9%; Score 58.8; DB 9; Length 996;
Best Local Similarity 33.1%; Pred. No. 0.088;
Matches 176; Conservative 80; Mismatches 276; Indels 0; Gaps 0;

QY      25 TTATATTGTATGACAAAGGATTAACCTAGTTATGATTAATTAATTAATTAATTAATTAATC 84
Db      465 YMCVHNNNNAAAAAAYAGAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 524
QY      85 AATTAAACGACATCACGCTCATAGACGGCTTGAGAGAAGATTAATATGATGTTCATCA 144
Db      525 AAAAAAAAAAAAAAAAAAARATMAAATAATAATAAAAAAAAAATWAAAWTWAHTMA 584
QY      145 ATATTTCAGATCATTAATGATAGGGGAATATCAAGCAATTAATAATGATTAATGCGATT 204
Db      585 TATTTAAAAAAAAAWTTTCWAAATTTWAAATTTWTAATTAATAATTAATTAATTAATTAAT 644
QY      205 AAAGTAAATAAGGAGAAATGATTCACCCCAATTTGAATGAGGTGGATGATTTCTTCTTTG 264

```

```

Db      645 AAAAAAAWAAAAANGTTTAAATWMAATCAWTAATAATAATAATAATAATAATAATAATA 704
QY      265 ACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 324
Db      705 AAAAAAAWAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 764
QY      325 AGTATGGAATAGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 384
Db      765 AAAAAAAWAAATTTTATATWAAWAAATWAAWAAATWAAWAAATWAAWAAATWAAWAAAT 824
QY      385 AGAGAAAGTCTGCTTTTCAAGAAATATTTTATCAGAGAATATTTATACATCCCTCTCTCC 444
Db      825 WAAAHWATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 884
QY      445 CTATCTCTTTTCTATTTATATGAGACATTTCTCTCAATCAATCTCTCTCTCTCTCTCTCT 504
Db      885 AAATWMTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 944
QY      505 AGAATATTCATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 556
Db      945 WATWTWMTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 996

RESULT 9
AG311015      2015 bp      DNA      linear      GSS 02-JUN-2004
LOCUS      Mus musculus molossinus DNA, clone: MSMg01-093J24.TJ, genomic survey
DEFINITION      sequence.
ACCESSION      AG311015
VERSION      AG311015.1 GI:47883969
KEYWORDS      GSS.
SOURCE      Mus musculus molossinus
ORGANISM      Mus musculus molossinus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      1
JOURNAL      Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
COMMENT      BAC end Sequences of Library MSMg01
      2 (bases 1 to 2015)
      Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
      Direct Submission
      Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
      and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
      1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
      (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
      Tel: 81-45-503-9111, Fax: 81-45-503-9170)
      Clones are derived from the mouse BAC library MSMg01. For BAC
      library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
      Tsukuba Institute, Bio Resource Center,
      The Institute of Physical and Chemical Research (RIKEN) 3-1-1
      Koyadai, Tsukuba, 305-0074 Japan
      Phone: 81-298-36-9189, fax: 81-298-36-9199
      e-mail: abe@rtc.riken.jp
      PRIMERS
      Sequencing : TJ
      LIBRARY      : pBACe3.6
      Vector      : EcoRI
      R.Site 1    : EcoRI
      R.Site 2    : EcoRI
      Location/Qualifiers
      1..2015
      /organism="Mus musculus molossinus"
      /mol_type="genomic DNA"
      /sub_species="molossinus"
      /db_xref="taxon:57486"
      /clone="MSMg01-093J24.TJ"
      /sex="male"
      /tissue_type="mixture of kidney and spleen"
      /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

```

Query Match 5.9%; Score 58.8; DB 9; Length 2015;
 Best Local Similarity 42.0%; Pred. No. 0.086;
 Matches 386; Conservative 1; Mismatches 532; Indels 1; Gaps 1;

QY 8 ACTTTGATTAAATTAATTAATATTTGATGACAAAGGATTAAACCTAGTTAATGATATA 67
 DB 234 AATTATATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 293
 QY 68 ACTTCAGATCTATATCAATTAACAGCAATCACGGTCATAGCAGCGTTTGAGAGAATTA 127
 DB 294 AATAAACCAAAAAAATATAGATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 353
 QY 128 AATGATGATGTCATTCATTAATTTCAAGATCTTAATGATAGGGAATATCAAGCAATAA 187
 DB 354 AATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 413
 QY 188 TAACGATAAATGGCATTAAGTAATAAGGAGATGATTCACCAATATTGAATGAGGTG 247
 DB 414 AATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 473
 QY 248 GATGATCTCTTTTGACAAATGATGAATGATGGGCAAACTACTAGATGTTGGACCCCTT 307
 DB 474 AATAATATAAATAAACCGTAAATAATATATAATAAATAAAAAAATAAAAAAATAATA 533
 QY 308 CTCGATCTAATGAATAAGATGATAGATAGATATCGAATCTCTTTGAAGAGGTAGT 367
 DB 534 TTATTTTATAAACAATATAAAGATTAATAAATAAATAAATAAAGAAAAAATAAAT 593
 QY 368 GATGTCCTTTTATCTAGAGAGAAAGTCGTCTTTTCAAGAAATATTTTTATCAGAGATAT 427
 DB 594 TTAGATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 653
 QY 428 TACATCCCTCTCTCCCTATCTCTTTTCTTTTATTTATATGGGACATTCCTCAATCATCC 487
 DB 654 AAAAAAATAAANAATAAATAATTTTTTTTTTATATTTTTTTTACAAAAAGNAAAAATAA 713
 QY 488 TAAAGATACATACACCAAGAAATATCAATAAATAATTTTTTGAATATCTTATTATAAAA 547
 DB 714 TAATAAAGATATAAGATAAGATACTATCTATAAGTATCTAAGATAAATAAATAAAGA 773
 QY 548 ACTAGCTGTTAGACTGACCTGGTGYATTGACTACTCGGTTAGAGCCCTGTGCTT 607
 DB 774 AAAAGTANCAACCCACACCTTTATTTTTTATTTATTTCTCTCCCTCCCTCCCTCCACC 833
 QY 608 TACTAATCGCTCGATTCATCATCTTCTACGATGCTGCTTCATGTCATCTTATGA 667
 DB 834 ACRAAAAAAATAAATAAATAAATACTTTCTTTTCAAAAAAATAAAAAAATAAAG 893
 QY 668 AAGCAGATTTTGACCCATACATAATATGACAAAATTTGCTTCCAAAGAAAAACATGGCTCT 727
 DB 894 AAAAAAGAGGGGGTAAATAATTTTATATCTAGAAATATATAAAAAAATAAAAAATA 953
 QY 728 TATAGTGAATATCGTTAGACTGTTATAGAAGATCTGAATTTATTTATAGAATAGTGT 787
 DB 954 TATAATAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1013
 QY 788 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 847
 DB 1014 AATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1072
 QY 848 ACTATATATCAAGGAATGGTGTGTTTTTCTTTTAAATATGGAATAAATTTGTGAATAG 907
 DB 1073 AATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1132
 QY 908 AAGATAGATCAATTAACAA 927
 DB 1133 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1152

RESULT 10
 CNS00DKY 928 bp DNA linear GSS 04-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
 DEFINITION

BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit fly), Genomic survey sequence.
 AL071865
 VERSION AL071865.1 GI:4948170
 GSS.
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 928)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammos in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 Location/Qualifiers
 1..928
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR27A24"
 /clone_lib="RPCI-98"
 /note="end : T7"

ORIGIN
 Query Match 5.9%; Score 58.8; DB 9; Length 928;
 Best Local Similarity 32.1%; Pred. No. 0.097;
 Matches 157; Conservative 67; Mismatches 258; Indels 7; Gaps 1;

QY 452 TTTTCTATTATATGGGACATTCCTCAATCAATCCTAAAGATACATACCAAGATAT 511
 DB 287 TTATTTCCAGAGACAACACAGACATTAAGATCCAAATATATATATATATATATAT 346
 QY 512 TCAATAAAATATTTTTTGAATATTCCTATTATAAAAACTAGCTGTTAGCACTCGACCTCG 571
 DB 347 AT 406
 QY 572 GTCGYATTGACTACTCGGTTAGAGCCCTGCTATTACTAATCGACCTCGATTACATCA 631
 DB 407 GGGCCNNCCNCCANANANNNCCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 466
 QY 632 CTTTCTACGATATCTGCTCATGTCATAATCTTAATGAAGACAGATTTTGACCCATCAATA 691
 DB 467 NNCNNNGGATTTTAT 526
 QY 692 ATATGACAAAATTTGCTTCCAAAGAAAAACATGCTCTTAT-----AGTGAATATCGTT 744
 DB 527 ANAATTTTTTTTTTTTTTAAATAATTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTT 586
 QY 745 AGACTGTTATAGAAGATCTGAATTTATTTATAGAATAGTGTGTTTTTTCTTTCTTTCT 804
 DB 587 TTATTTAAAWTANWATTTTATATATATATATATATATATATATATATATATATAT 646
 QY 805 ATATCTAAGGAGTAAAGCAACCATGAATAGAAAAAGGCTTAGTAACATATATATATAT 864
 DB 647 WTTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 706

ORIGIN

```

ORIGIN
  (http://www.chori.org/bacpac/ordering_information.htm).

Query Match      5.8%; Score 58.4; DB 9; Length 1335;
Best Local Similarity 40.0%; Pred. No. 0.11;
Matches 214; Conservative 1; Mismatches 319; Indels 1; Gaps 1;

QY 15 ATTAAATTAATAATTTGTATGACAAAGGATTAACCTAGTCTTAATGATATAACTTTCAG 74
Db 1122 ANAAAAATTAATTTTTTATATAATAAAAAAATAAAAAATANAATAAATAAAAAATATAA 1063
QY 75 ATCTATATCAATTAACAGCAATCACGCTATAGCAGCGTTGACAGAGATTAATGTA 134
Db 1062 AAATANATTTNTATTATAAAAAAANTATAATTTTATATAAAAAAATAAATATNTNAA 1003
QY 135 TGTCAATTCATATTTCAAGATCATTAATGATGGGAATATCAAGCAATAAATAACGAT 194
Db 1002 NNATATNAANNNTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATTT 943
QY 195 AAATGGCATTAAGTAATAAGGAGATGATTCACCCATATTCGAATGA-GGTGGATGAT 253
Db 942 ATTTTAAATNNNAATTTNATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 883
QY 254 TCTTCTTTTTCACATGATGATGGCAATAGTGGCAATAGTGGCAATGTTGGACCTTCTCGGA 313
Db 882 NATTAATTAATTTATAATAAATAATTTANNTATANAATAAATAAATAAATAAATAAATAA 823
QY 314 TCTAATGAAAGATGATGGAATAGTAGAATAATCGAATCTCTTTAGAAAGGTAGTGTGT 373
Db 822 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 763
QY 374 CTTTTATCTAGAGAGAGTCTGCTTTTCAAGATATTTTATCAGAGATATTCATC 433
Db 762 TATTTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 703
QY 434 CCCCTCTCTCCTATCTCTTTTCTATTATATGAGGACATTCCTCAATCAATCCTAAAG 493
Db 702 TNNANNTATATATNTNTNTNTATATATATATATATATATATATATATATATATATATA 643
QY 494 TACATACCAAGAAATATTCATATAAATAATTTTTTTGAAATTCCTATATAAATAA 549
Db 642 TATATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 588

RESULT 13
CNS003BD 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC08K08 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL064091
AL064091.1 GI:4941847
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial

```

```

EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
PI and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
      Location/Qualifiers
          source
              1..1101
              /organism="Drosophila melanogaster"
              /mol_type="genomic DNA"
              /db_xref="taxon:7227"
              /clone="BACH08K08"
              /clone_lib="RPCI-98"
              /note="end : TET3"

ORIGIN
Query Match      5.8%; Score 58; DB 9; Length 1101;
Best Local Similarity 35.9%; Pred. No. 0.13;
Matches 207; Conservative 58; Mismatches 312; Indels 0; Gaps 0;

QY 238 GAATGAGGTGGATGATCTCTTTTGGACAATGATGATGGCAATATCTAGAAATCT 297
Db 1087 GGATAGWTGKDTKWTWTWTWATAATNTADTTKTTTWTATATWATAKAWTTTTTWT 1028
QY 298 TGGGACCCCTTCGGAATCTAATGAAAGATGGAATAGTAGATAATCGAATCTCTTA 357
Db 1027 RWTWTTRTWTWNTWNTWNTTATTTTATTTTWTATTAATATWATWATAAATAA 968
QY 358 GAAAGGTAGTGTGTTCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAAATATTTTAT 417
Db 967 TATTTWTWTAATWRWTTTTATARAAGAAWAAWTATWATTTTATTTTATTTTATTTT 908
QY 418 CAGAGAAATATACATCCCTCTCTCCCTATCTCTTTTCTATTTATATATGGACATCT 477
Db 907 WTTATWTTTATTTTATTTAAAWTATWATAWATTTATTTATTTTATTTTATTTTAT 848
QY 478 CAATCAATCCTAAAGTACATACACCAAGAAATATCAATAAATAATTTTGTGAATATTC 537
Db 847 TWTANATWTTTAAKAWATTTAAWATTAATAAATAAATAAATAAATAAATAAATAA 788
QY 538 TATTATAAAGCTAGCTGTGTAGACATCGACCTCGGTGCGTATTTGACTACTCGGTACGAG 597
Db 787 WATWATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 728
QY 598 CCCTGTCATTTACTAATCGACCTCGATTCATCATCTCTACGATACGCTTCATGTCGA 657
Db 727 TTTAWATTTATAATTTTTTTTATATTTTTTTTTTTTTTTTTTTTTTTTAAATTTNWTAA 668
QY 658 ATCTTAATGAAAGCAGATTTTGACCCCATACAAATAATGACAAATTTGCTTCCAAAGAAA 717
Db 667 ATAAAAATTTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 608
QY 718 ACATGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGAATTTATTATA 777
Db 607 AAATKWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATKATAAATAAATA 548
QY 778 AGAATAGTCTTTTTTTTCTTTCTTTTCTATCTAATCTAAGG 814
Db 547 AAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTAAGG 511

CNS014J2 1025 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC
BACH11L1 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL104216
AL104216.1 GI:5615827
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source

Location/Qualifiers
1..1025
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN1101"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

ORIGIN

Query Match 5.7%; Score 57.4; DB 9; Length 1025;
Best Local Similarity 37.2%; Pred. No. 0.17;
Matches 121; Conservative 50; Mismatches 154; Indels 0; Gaps 0;

QY 622 GATTACATCCTTCTAGCACTCTCTCATGTCAGATCTTAATGAAAGCAGATTTTGCAC 681
DB 744 GTATWATAAATATTKKTGTGTARAWAAATATTAATAWAYGAKTGRABAA 685
QY 682 CCAATCAATATATGACAAATATGCTTCCAAAGAAACATGGCTCTTATAGTGAATATC 741
DB 684 AAARAAAAAATGTRAWATATWATTTKATTTGTGAATWKTGTTKAWAAAAA 625
QY 742 GTTAGACGTATAGCAAGATCTGAATTTATTTAAGAATAGTGTCTTTCTTTCTT 801
DB 624 AAATAKWTAAWKTAAAHNDGTTKTTTAAWAAKGGRTGKAKGTTRDWW 565
QY 802 TTCAATCTAAGAGTAAAGCAACATGAATAGAAAAGCTTAGTAATATATCAAG 861
DB 564 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 505
QY 862 GAATGCTGTTTTCTTTTAAATATGATAAATAATTTGTAATATAGAGATTAGATCAAT 921
DB 504 KTGCTKTKTGCTAGTAAGAKTAAATWAAAGTTGKAKWATATTAATAAATAAATAA 445
QY 922 TAACAAAGGTATGCTGGAGTGGTA 946
DB 444 AAAAAKWTGCTGKGAATTAAT 420

RESULT 15
CNS0152Z
LOCUS

Drosophila melanogaster genome survey sequence T7 end of BAC
BACN15K05 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION
VERSION
KEYWORDS

AL106121.1 GI:5620267

SOURCE

Drosophila melanogaster (fruit fly)

ORGANISM

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 1201)

Direct Submission

JOURNAL

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source

Location/Qualifiers
1..1201
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15K05"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

ORIGIN

Query Match 5.7%; Score 57.4; DB 9; Length 1201;
Best Local Similarity 32.4%; Pred. No. 0.17;
Matches 222; Conservative 102; Mismatches 355; Indels 6; Gaps 1;

QY 250 TGATTCCTCTTTTGACATGATGATGCGGCAAACTAGAAATGTTGGACCTTCT 309
DB 476 TSTATTCTTATTTGTSSTNNNTAANTNTTATSNSTNANWSTNNNNNTTTT 535
QY 310 CGGATCTAATGAAAAAGATGGAATAGTAGAATATCGAATCTCTTAGAAGGTAGTA 369
DB 536 TTTTAAAAAANNANNAVAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 595
QY 370 TTGCTCTTTATCTAGAGAGAAAGTCTGCTTTCAAAGATATTTTATCAGAGATATTA 429
DB 596 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 655
QY 430 CATCCCTCTCTCCCTATCTCTTTTCTATTATATGGAATCTCTCAATCAATCTTA 489
DB 656 AAWCATCCACTTWTATTATTTTCTTTTATATTTTATCCATCMCCATTTTWT 715
QY 490 AAGATACATACACCAAGATATCAATAAATAATTTTGAATATCTTATATAA 549
DB 716 TTTTTCACAWCAAAWAAACCCCTATATTTTATTTTATTTTATTTTATTTT 775
QY 550 TAGCTCTTAGCACTCGACCTCGTATGACTACTCGGTACGAGCCCTGTCATTTA 609
DB 776 HWAHAHAHAACMAACVTTTTTCTTTTACAMHYCTTTAHTCTTTTCTTCTCC 835
QY 610 CTAATCGACCTCG-----ATTACATCACTTCTTAGGATCTGCTTCATGCAATCTTA 663
DB 836 CTAAWAAATCCCTCTCAATTTTWTMTTCHAWCCCTHTTTCYCYCCMYHHCHCTY 895
QY 664 ATGAAGACAGATTTTGACCCATACATAATATGACAAATTCCTTCCAAAGAAACATGG 723
DB 896 YCTWAABTTTAATTTTCTTTTTSSTAMTSYTMGCATTTTCHMYCYCYWAWHHH 955
QY 724 CTCCTTAGTGAATAATCGTTAGATCTGTTATAGAAGATCTGAATTTTATAGAATA 783
DB 956 HHHHHHHYHYMMWAAWTTTTTTTATTTTTCACAGSYTTAWAAMWMTTWT 1015
QY 784 GTGTTTTTTTCTTTTCTTTTCTATCTAGGAGTAGTAAGCAACCATGAATAGAAAAGCTT 843
DB 1016 TWAAMWTTTWTTTTCTTTTCTTTTAAAAAATTAATAAAAAAATAAAAAA 1075
QY 844 AGTAACATATATCAAGGAATGGTGTCTTTTCTTTTAAATATGGAATAAATTTGTAAT 903
DB 1076 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1135
QY 904 ATGAAGATAGATCAATTAACAAA 928
DB 1136 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1160

Search completed: November 11, 2004, 12:05:44
Job time : 3112.21 secs

This Page Blank (uspto)

Result No.	Score	Query		DB ID	Description
		Match	Length		
1	60.4	6.0	1141	4	US-09-806-708B-22
2	52	5.2	1141	4	US-09-806-708B-22
3	50.4	5.0	1055	4	US-09-806-708B-23
4	49.2	4.9	5360	4	US-10-204-708-66
5	48.8	4.9	832	4	US-09-621-976-2813
6	47.2	4.7	1055	4	US-09-806-708B-23
7	46.4	4.6	1122	4	US-09-248-798A-729
8	46	4.6	1140	4	US-09-270-767-29120
9	46	4.6	7218	4	US-09-270-767-13204
10	45.8	4.6	2396	1	US-08-233-463-14
11	44.6	4.5	319608	4	US-09-539-333D-1
12	44.6	4.5	319608	4	US-09-679-403-1
13	43.8	4.4	4285	3	US-09-410-464-1
14	43.6	4.4	392	4	US-09-270-767-29271
15	43.6	4.4	2031	4	US-09-270-767-13327
16	43.4	4.3	6113	4	US-10-204-708-13
17	43	4.3	933	4	US-09-248-796A-9931
18	42.6	4.3	3867	4	US-09-366-715-5
19	42.2	4.2	1924	2	US-08-487-828B-13
20	42	4.2	539	4	US-09-270-767-25193
21	42	4.2	643	4	US-09-270-767-30522
22	42	4.2	1188	4	US-09-270-767-9895
23	42	4.2	1842	4	US-09-270-767-14367
24	41.8	4.2	1316	4	US-09-270-767-2481
25	41.8	4.2	1316	4	US-09-270-767-17763
26	41.8	4.2	640681	4	US-09-790-988-1
27	41.6	4.2	2238	4	US-09-799-024-34
28	41.6	4.2	2238	4	US-09-799-024-34
29	41.6	4.2	2238	4	US-09-799-024-34
30	41.6	4.2	2238	4	US-09-799-024-34
31	41.6	4.2	2238	4	US-09-799-024-34
32	41.6	4.2	2238	4	US-09-799-024-34
33	41.6	4.2	2238	4	US-09-799-024-34
34	41.6	4.2	2238	4	US-09-799-024-34
35	41.6	4.2	2238	4	US-09-799-024-34
36	41.6	4.2	2238	4	US-09-799-024-34
37	41.6	4.2	2238	4	US-09-799-024-34
38	41.6	4.2	2238	4	US-09-799-024-34
39	41.6	4.2	2238	4	US-09-799-024-34
40	41.6	4.2	2238	4	US-09-799-024-34
41	41.6	4.2	2238	4	US-09-799-024-34
42	41.6	4.2	2238	4	US-09-799-024-34
43	41.6	4.2	2238	4	US-09-799-024-34
44	41.6	4.2	2238	4	US-09-799-024-34
45	41.6	4.2	2238	4	US-09-799-024-34
46	41.6	4.2	2238	4	US-09-799-024-34
47	41.6	4.2	2238	4	US-09-799-024-34
48	41.6	4.2	2238	4	US-09-799-024-34
49	41.6	4.2	2238	4	US-09-799-024-34
50	41.6	4.2	2238	4	US-09-799-024-34
51	41.6	4.2	2238	4	US-09-799-024-34
52	41.6	4.2	2238	4	US-09-799-024-34
53	41.6	4.2	2238	4	US-09-799-024-34
54	41.6	4.2	2238	4	US-09-799-024-34
55	41.6	4.2	2238	4	US-09-799-024-34
56	41.6	4.2			

Db 359 KARHBAWDMVHSAWKHANAHAHSEKKBWBYKRTKTMNNNNNGTTTMMKRMNAWYKMD 418
 QY 371 TGCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAATAATTTTATACAGAGAAATATAC 430
 Db 419 MDWBTNNNNNGRTYGTGKWKWYKWKANNCKWRADHDKTCTHNTTWWKXKY 478
 QY 431 ATCCCTCTCTCCCTATCTCTTTTCTATTATATAGGACATTCCTCAATCAATCTCA 490
 Db 479 WNNCVKSMINGKSHRBAAAVYTWYMMWRRAHANNNNDYWKACTWYKBYCVSKWN 538
 QY 491 AAGTACATACCAAGATATTCATAAATATTTTTCGAATATCTATTATAAAACT 550
 Db 539 NYAATYKSSWNTSYRKYRWKTNNSWRSDTRSMGRANNYARABHYKYKWTIRWBWSH 598
 QY 551 AGCTGTAGTACCTCGACTCGGTCTGCTTCTGCTTACGATCTGCTTACGACCTGCTTAC 610
 Db 599 TWBHRAGAAYHWMYBAMVYBAKCHWKAWYKAKYVAGAGSNNNNNNNNNNNNNNATCA 658
 QY 611 TAATCGACCTCGATACATCACTTCTGCTTACGATCTGCTTCAATCAATCTTAAAG 670
 Db 659 RDDYASRWYMANAKWYKYYKBAANNAYYTHANNWGCWNNATDTTTRTWKNNNNNAG 718
 QY 671 CAGATTTTGCACCCATACATATGACAAATATGCAAAATGCTTCCAAAGAAACATGG----- 723
 Db 719 TWKNNNNNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAK 778
 QY 724 CTCCTTATAGTAAATATCGTGTAGTCTGCTTATAGAAATCTGATTTTATTATAAGATA 783
 Db 779 KYTTNNNTYRGVVTNTAARDGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 838
 QY 784 GTGTTTTTCTTTCTTTTCTTATCTAAGAGAGTAAAGCAACATGAATAGAAAGGCTT 843
 Db 839 NNNNNAYAWNTKYYITTDWRBAYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 898
 QY 844 AGTAACATATATCAAGAAATGCTTTTCTTTTAAATGATGATAAAATTTGGAAT 903
 Db 899 DATKNNATTYRGTAWRNN 958
 QY 904 ATAG 907
 Db 959 TTKG 962

RESULT 2
 US-09-806-708B-22/c
 ; Sequence 22, Application US/09806708B
 ; Patent No. 6784342
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
 ; FILE REFERENCE: 4810-58741
 ; CURRENT APPLICATION NUMBER: US/09/806,708B
 ; CURRENT FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: US 60/147,133
 ; PRIOR FILING DATE: 1999-08-04
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 1141
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (1)..(1141)
 ; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
 US-09-806-708B-22

Query Match 5.2%; Score 52; DB 4; Length 1141;
 Best Local Similarity 7.9%; Pred. No. 0.0019;
 Matches 46; Conservative 260; Mismatches 274; Indels 0; Gaps 0;
 QY 246 TGGATGATCTCTTTTGTGACATGATGAATGATGGGCAATATAGATGTTGGGACCC 305

Db 633 TRMTVRMTKGDGMVVRKKVKWRDCTTCTYVDVWADSWVWVYANMRCRDYTYTRNTYC 574
 QY 306 TTCTCGGATCTAATGAAAAAGTATGGAATAGTAAATCAATCTCTTTAGAAAGGTA 365
 Db 573 KSYAHSYVWVSNAMWYRRYSARNWSSMARWTRNNNNWMSGBVRMAGTMMWRHNNNN 514
 QY 366 GTGATGCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAATAATTTTATCAGAGAT 425
 Db 513 TDTTRYWKKRAREBTITTYDSMCKNAKSMRGNWRAKQWMAANNDAGMDHWTYMGNN 454
 QY 426 ATTACATCCCTCAAT 485
 Db 453 NTWMBRAKWMWMAWCRRAAYCCNNNNNRACVWHKHKWRWTKYMWKACANNNNBKWY 394
 QY 486 CTTAAAGTACATACCAAGAAATATCAATAAAATATTTTTCGAATATCTTATATAA 545
 Db 393 MRVAVWYSRDTTNTDMMWTSDBWHWTVDYTMRAWNNNNNNNNNNNNNNNNNNNNNN 334
 QY 546 AAATAGCTGTAGCACTCGACTCGGTCTGCTTATGACTACTCGGTACGAGCCTGTCA 605
 Db 333 HNTHTCTGNTWGSAYENMAWMSWAAGASNBVTYNWCRTWYMGKTNNNNNNKAWY 274
 QY 606 TTTACTAATCGACCTCGATATCACTTTCTGATATCTGCTTCAATCTCAATCTTAA 665
 Db 273 RUTVAVCMNNRYYYDTAVMTBKRYKYCYBYWYBYMTMGKHHWBRABHRSNNWV 214
 QY 666 GAAAGCAGATTTGACCCATACATAATATGACAAATGCTTCCAAAGAAACATGGCT 725
 Db 213 KCRNYKYSVWHYHARVKBWABAVGCCNNNNKDRMAHHWCAATNNNNNNNNNNNNNN 154
 QY 726 CTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGAAATTTATTATAAGAAT 785
 Db 153 GAAWNTNKTABRDBHBAHVXYWYRVDYWCAMCMWNAKAKVRTAMCHWYTDYVSA 94
 QY 786 GTTTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 825
 Db 93 NNTGYRMMWRCMMWYSMNNRRYMRGRKYTWAMYSMS 54

RESULT 3
 US-09-806-708B-23
 ; Sequence 23, Application US/09806708B
 ; Patent No. 6784342
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
 ; FILE REFERENCE: 4810-58741
 ; CURRENT APPLICATION NUMBER: US/09/806,708B
 ; CURRENT FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: US 60/147,133
 ; PRIOR FILING DATE: 1999-08-04
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 23
 ; LENGTH: 1055
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (1)..(1055)
 ; OTHER INFORMATION: consensus sequence of A.t. and L.a. FAE1 promoters
 US-09-806-708B-23

Query Match 5.0%; Score 50.4; DB 4; Length 1055;
 Best Local Similarity 20.6%; Pred. No. 0.0045;
 Matches 181; Conservative 188; Mismatches 504; Indels 5; Gaps 1;
 QY 13 TCATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72
 Db 3 TSKAAWAAARMYAKYAGWNTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 62
 QY 73 AGATCTATAATCAATTAACAGCAATCAAGTCATAGCAGCGTTGAGAGAGATTAATGT 132

Db 63 TKWGTTCWAWTWRWAAWKKTKWMCSTAMNNWTTWCTARKWRTGTRWTKNNATGT 122
Qy 133 GATGTYCATTCATTAATTTCAAGATCAATTAATGATAGGGAATATCAAGCAATATAAAG 192
Db 123 RWTGYWNTNNNGSNTWARRYKTRWMCYTAWMYGASWAGNASTRTTYTWYRWKMK 182
Qy 193 ATAAATGGCAATTAAGTAATAAGGAGAATGATTCAACCAATATTTGAATGAGGTGGATGA 252
Db 183 RKSARATRCGRARYMRAWWTWARPRTGWTKAWAYAAWTTWNNNNNAKAACKAATTWGR 242
Qy 253 TTCTTCTTTTGAACAATGATGATGGCAAAATCACTAGAATGTTGGGACCCCTTCCTCGG 312
Db 243 AKSNCTCTTAGTTTTRATCCWATTCGAGWATKKWTKTKTSAAMGWTWNNNNNNNTTTK 302
Qy 313 ATCTAATGAAAAGTAGTAGATAGATAATCGAATCTCTTTAGAAAGGTAGTAGTG 372
Db 303 AAMTAARAWWWSATTTWAAAATTSRTWYGRKTANNNGTTCWTRWATWKKMK 362
Qy 373 TCTTTTCTAGAGAGAAAGTCTCTTTTCAAGAAATPATTTTATACAGAGAATATTACAT 432
Db 363 TKGTTWNNNGRRTYTCGTTTCKMATTTTAKANNCTTAAMKWKTCCTMNTTAAKATTWAT 422
Qy 433 CCCCTCTCTCCCTATCTCTTTTCTATTTATATGAGACATTCTCTCAATCAATCCTAAAA 492
Db 423 CYWKSMTNGTSYRYAAARYTWYAWWTR-----RYAYANNNTKTWKAACCTWTKRCCTT 477
Qy 493 GTACATACACCAAGAAATATCAATAAATAATTTTTTGAATATTTCTATTATAAAAACTAG 552
Db 478 ANNTAAYTKSSANCTSRTRWKNCRAGSKTASMGRAVARAYWTKWNTAWYCWWT 537
Qy 553 CTGTTAGACTCGACCTCGGTGGVATGACTACTCGGTACGAGCCCTGCTCATTTACTA 612
Db 538 WYVRAGAAMTMYMTSATCYCAATAATTAGTCAGAGGSTAKGNNNNNNNNCCATCAR 597
Qy 613 ATCGACCTCGATTACATCACCTTCTACGATAGTCTTCATGTCAAATCTTAATGAAGCA 672
Db 598 WKCTAASACMANATTCYCYAANNATYWANATCGWNAKTATATWNNNNNNNGTW 657
Qy 673 GATTTTGACCATACAAATAATATGACAAAATTTGTTCCAAAGAAAACATCGCTTTATAG 732
Db 658 TNNNNNAKVASATWYAAAMTAATKYARTANTAMAGAYARAAAYTTRTANNGACTTT 717
Qy 733 TGAATATCTGTAGACTGTTATAGAAAGATCTGNAATTTTATAGAAATAGTGTTTT 792
Db 718 TTTNNTTGGWRTNTAAARGWNNNNNNNNNNNNNNNGACWARTTTATANCSTNNNNNNN 777
Qy 793 TCTTTTCTTTTCAATCTAGGAGTAAAGCAACCATGAATAGAAAAGCGCTTAGTAATAT 852
Db 778 NAYATTTNTATTTTWTTRKANNNNNNNNNNAAYYGAAAKNNNTTMCWTKAWKAWATGA 837
Qy 853 ATATCAAGGAATCGGTGTTTTTTCTTTTAAATATGGATA 890
Db 838 ATTTNAGTATATNNNNNNATATTTTYYKAATNGKACTA 875

RESULT 4
US-10-204-708-66
; Sequence 66, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8

; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 66
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-66
Query Match 4.9%; Score 49.2; DB 4; Length 5360;
Best Local Similarity 52.4%; Pred. No. 0.014; 98; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 0
Qy 684 ATACAATAATGACAAAATTCCTTCCAAAGAAAACATCGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAAAAATTTGGATATAGTAGTATTGGTTATTATTAAAGATGATA 1852
Qy 744 TAGACTGTTATAGAAGATCTGAATTTATTTATAAGATAGTGTGTTTTCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAAGTTTTATGTAAAGATATATTTTATTATTAGTTA 1912
Qy 804 CATATCTAAGGAGTAAAGCAACCATCAATAGAAAAAGGCTTAGTAATATATATCAAGGA 863
Db 1913 GTTTTAAAAACGTTAGGAAGAGTATATATTAATGAAGTATTTTATGGGTAATATTA 1972
Qy 864 ATGGTGTGTTTTCTTTTAAATATGGAT 889
Db 1973 AAGTTGTTTTTATTTTGTAAATTCGAAT 1998
RESULT 5
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813
Query Match 4.9%; Score 48.8; DB 4; Length 832;
Best Local Similarity 10.8%; Pred. No. 0.01;
Matches 29; Conservative 136; Mismatches 103; Indels 0; Gaps 0;
Qy 721 TGGCTCTTATAGTGAATATCGTTAGACTGTATAGAAGATCTGAATTTATTTATAAGA 780
Db 1 YAWYWKYTTTAKCWTWKWSWSYWKYTKYTWYWRKWKKAWKWKYTKTWYWW 60
Qy 781 ATAGTGTGTTTTTCTTTTCTTATCATATCAAGCAACCAATGAATAGAAAAGG 840
Db 61 RYAMGTYKKKAMCRKTKKKKKKGYMMWYWGWSRYAMWTRTWGTVAYYRSWYWWR 120
Qy 841 CTTAGTAATATATCAAGGAATGCTGTTTTTCTTTAAATATGGAATAAAATTTGTG 900
Db 121 YRCWKKAAYRKTCTYSSKGWTTWKRKCAWTTWWKKYTYWAAATRYWMMCMCTKWRAS 180

Db 302 CTATTGAAGCCTCTTAATTATTATTGAAGCTGCARATCCGGGATCACCAGATGGATAA 361

QY 198 TGGCATTAAGTAATAAAGAGGATGATTCACCCAA 233

Db 362 TCGCATTAAAGTACATAAATAAATAAATAAACCA 397

RESULT 9

US-09-270-767-13204

; Sequence 13204, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 13204

; LENGTH: 2396

; TYPE: DNA

; ORGANISM: *Drosophila melanogaster*

US-09-270-767-13204

Query Match 4.6%; Score 46; DB 4; Length 2396;

Best Local Similarity 50.5%; Pred. No. 0.066;

Matches 109; Conservative 1; Mismatches 106; Indels 0; Gaps 0;

QY 18 AATTAATTAATATTGATGCAAGAGGATTAACCTAGTTAATGATAATTAATTCAGATC 77

Db 182 AACATAAATTGTAGTTCTTTGAAAGAGATTAACCAACACTGGGAATTAGAAAATAAAC 241

QY 78 TATAATCAATTAACAGCAATCACGGTCATAGCAGCGTTGAGAGAGGATTAATGTGATGT 137

Db 242 GAAAAATCTTTAGAAACTTAATAAGGAAACTGAAAGAAATCGAGGCATATAT 301

QY 138 YCATTCATATTTCAAGATCATTAATGATAGGGGAATATCAAGCAATAAATAACGATAA 197

Db 302 CTATTGAAGCCTCTTAATTATTATTGAAGCTGCARATCCGGGATCACCAGATGGATAA 361

QY 198 TGGCATTAAAGTAATAAAGGAGATGATTCACCCAA 233

Db 362 TCGCATTAAAGTACATAAATAAATAAATAAACCA 397

RESULT 10

US-08-232-463-14/c

; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313

; FILING DATE:

; APPLICATION NUMBER: EP 91 114 300.6

; FILING DATE: 26-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 30472/114 INMU

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEFAX: (703)683-4109

; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7218 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; CLONE: pTZ9pt-F1s

US-08-232-463-14

Query Match 4.6%; Score 45.8; DB 1; Length 7218;

Best Local Similarity 3.7%; Pred. No. 0.1;

Matches 14; Conservative 210; Mismatches 157; Indels 0; Gaps 0;

QY 12 TTGATTAATAATTAATAATTTGATGACAAAGGATTAACCTAGTTAATGATAAATACTT 71

Db 1459 TTAAGAGATAGAGAATTTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1400

QY 72 CAGATCTAATCAATTAACAGCAATCAGGTCATAGCAGCGTTGAGAGAGATTAATG 131

Db 1399 RRR 1340

QY 132 TGATGTTCATCAATTTCAAGATCATTAATGATAGGGGAATATCAAGCAATAATAAC 191

Db 1339 RRR 1280

QY 192 GATAAATGSCATTAAAGTAATAGAGAGATTAATCACCATAATGATGAGGTGGATG 251

Db 1279 RRR 1220

QY 252 ATTCTCTTTTGACAATGATGAATGATGGCAATAGTGGACCTTCCTCG 311

Db 1219 RRR 1160

QY 312 GATCTAATGAAAAAGATGGAATAGTAGAATAATCGAATCTCTTTAGAAAGTAGTGT 371

Db 1159 RRR 1100

QY 372 GTCTTTTCTAGAGAGAAAG 392

Db 1099 RRR 1079

RESULT 11

US-09-539-333D-1

; Sequence 1, Application US/09539333D

; Patent No. 6476208

; GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Chumakov, Ilya

; APPLICANT: Bouguetret, Lydie

; APPLICANT: Bihain, Bernard

; APPLICANT: Essicou, Laurent

; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS

; FILE REFERENCE: GENSET.047AUS

; CURRENT APPLICATION NUMBER: US/09/539,333D

; CURRENT FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: US 60/126,903

; PRIOR FILING DATE: 1999-03-30

; PRIOR APPLICATION NUMBER: US 60/131,971

```

; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 31..1107
; OTHER INFORMATION: 5'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1108..1289
; OTHER INFORMATION: exon A g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14877..14920
; OTHER INFORMATION: exon B g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18778..18862
; OTHER INFORMATION: exon Bbis g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25593..25740
; OTHER INFORMATION: exon C g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29388..29502
; OTHER INFORMATION: exon D g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29967..30282
; OTHER INFORMATION: exon E g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 64666..64812
; OTHER INFORMATION: exon F g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 65505..65853
; OTHER INFORMATION: exon G g35018 gene
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 65854..67854
; OTHER INFORMATION: 3'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 94124..94964
; OTHER INFORMATION: exon g35017
; FEATURE:
; NAME/KEY: exon
; LOCATION: 201188..201234
; OTHER INFORMATION: exon S g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 214676..214793
; OTHER INFORMATION: exon T g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215702..215746
; OTHER INFORMATION: exon U g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 216836..216915
; OTHER INFORMATION: exon V g35030 gene
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 213818..215818
; OTHER INFORMATION: 3'regulatory region g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215819..215941
; OTHER INFORMATION: exon R complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215819..215975
; OTHER INFORMATION: exon Rbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 216661..216952
; OTHER INFORMATION: exon Qbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 217027..217061
; OTHER INFORMATION: exon Q1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 229647..229742
; OTHER INFORMATION: exon X complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 230408..230721
; OTHER INFORMATION: exon P complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231272..231412
; OTHER INFORMATION: exon Obis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231787..231880
; OTHER INFORMATION: exon O2 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231870..231879
; OTHER INFORMATION: exon O1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 234174..234321
; OTHER INFORMATION: exon O complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 237406..237428
; OTHER INFORMATION: exon Nbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 239719..239807
; OTHER INFORMATION: exon N2 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 239719..239853
; OTHER INFORMATION: exon N complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240569
; OTHER INFORMATION: exon M117 complement g34872 gene
; FEATURE:

```

```

1 NAME/KEY: exon
2 LOCATION: 240528..240596
3 OTHER INFORMATION: exon M1090 complement g34872 gene
4 FEATURE:
5 NAME/KEY: exon
6 LOCATION: 240528..240617
7 OTHER INFORMATION: exon M1069 complement g34872 gene
8 FEATURE:
9 NAME/KEY: exon
10 LOCATION: 240528..240644
11 OTHER INFORMATION: exon M82 complement g34872 gene
12 FEATURE:
13 NAME/KEY: exon
14 LOCATION: 240528..240824
15 OTHER INFORMATION: exon M862 complement g34872 gene
16 FEATURE:
17 NAME/KEY: exon
18 LOCATION: 240528..240994
19 OTHER INFORMATION: exon M692 complement g34872 gene
20 FEATURE:
21 NAME/KEY: exon
22 LOCATION: 240528..241685
23 OTHER INFORMATION: exon M1 complement g34872 gene
24 FEATURE:
25 NAME/KEY: exon
26 LOCATION: 240800..240993
27 OTHER INFORMATION: exon M51 complement g34872 gene
28 FEATURE:
29 NAME/KEY: misc_feature
30 LOCATION: 241686..243685
31 OTHER INFORMATION: 5'regulatory region g34872 gene
32 FEATURE:
33 NAME/KEY: misc_feature
34 LOCATION: 290652..292652
35 OTHER INFORMATION: 3'regulatory region g34665 gene
36 FEATURE:
37 NAME/KEY: exon
38 LOCATION: 292653..292841

```

```
Query Match      4.5%; Score 44.6; DB 4; Length 319608;
Best Local Similarity 49.8%; Pred. No. 0.6;
Matches 113; Conservative 0; Mismatches 114; Indels 0;
```

QY	117	AGAGAAGATTAAATGTGATGTCATTCAATATTTCAAGATCATTAAATGATAGGGGTAAT	176
Db	104809	ACATAATTTTAAATTTAGAAAATTATGACAAAAC	104868
QY	177	CAAGCAATAAATAACGATAAATGGCATTTAAAGTAAATAAGGAGAAATGATTCACCCCAAT	236
Db	104869	CGATACTTTAAACCACTAGTAAATGTTAAGTCTCGCATTATGCA	104928
QY	237	TGAATGAGGTGGATGATTCCTCTCTTTTGACAATGATGAATGATGGGCAATAC	296
Db	104929	AGGATGACATGGCTCAGCGTTGTTTACAGAAATAGCACATGATAGCTGTTCCATAAATAC	104988
QY	297	TTGGGACCCCTTCTCGGATCTTAATGAAAAAGTATGGAAATAGTAGATA	343
Db	104989	TTGTTATATTGCTTGGGAGAAATGAGGCCCATATATTAATTCGAATA	105035

RESULT 12

```

; Sequence 1, Application US/09679409
; Patent No. 6585316
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Essicux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: 53.US15.C1P
; CURRENT APPLICATION NUMBER: US/09/679,409

```

```
OTHER INFORMATION: 99-31960-363 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 107281
OTHER INFORMATION: 99-24656-260 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 160640
OTHER INFORMATION: 99-24639-163 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 160876
OTHER INFORMATION: 99-24634-108 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 168974
OTHER INFORMATION: 99-7652-162 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 170810
OTHER INFORMATION: 99-16100-147 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 173358
OTHER INFORMATION: 99-5862-167 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 189957
OTHER INFORMATION: 99-5919-215 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 197163
OTHER INFORMATION: 99-24658-410 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 200778
OTHER INFORMATION: 8-303-235 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 202651
OTHER INFORMATION: 8-300-221 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 202679
OTHER INFORMATION: 8-300-193 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 203378
OTHER INFORMATION: 8-299-128 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 204138
OTHER INFORMATION: 8-296-213 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 204605
OTHER INFORMATION: 8-252-190 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 204934
OTHER INFORMATION: 99-24644-194 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 205206
OTHER INFORMATION: 8-295-248 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 205329
OTHER INFORMATION: 8-295-125 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 206054
OTHER INFORMATION: 8-293-130 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 206545
OTHER INFORMATION: 8-292-198 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 207313
OTHER INFORMATION: 8-251-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208285
OTHER INFORMATION: 8-289-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208960
OTHER INFORMATION: 8-287-249 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 209123
OTHER INFORMATION: 8-287-86 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 209631
OTHER INFORMATION: 8-285-319 : polymorphic base A or G
```

```
NAME/KEY: allele
LOCATION: 210361
OTHER INFORMATION: 8-283-278 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210463
OTHER INFORMATION: 8-283-176 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210486
OTHER INFORMATION: 8-283-153 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210583
OTHER INFORMATION: 8-283-56 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 210879
OTHER INFORMATION: 8-282-345 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210964
OTHER INFORMATION: 8-282-260 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 210979
OTHER INFORMATION: 8-282-245 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 211050
OTHER INFORMATION: 8-282-174 : variable motif AAAGG or GAAGGAAGGAAGGAAGA
NAME/KEY: allele
LOCATION: 211132
OTHER INFORMATION: 8-282-92 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 211247
OTHER INFORMATION: 8-281-367 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211315
OTHER INFORMATION: 8-281-299 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211366
OTHER INFORMATION: 8-281-248 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 212520
OTHER INFORMATION: 8-279-197 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 212821
OTHER INFORMATION: 8-278-289 : polymorphic base C or T
NAME/KEY: allele
```

```
Query Match 4.5%; Score 44.6; DB 4; Length 319608;
Best Local Similarity 49.8%; Pred. No. 0.6;
Matches 113; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 117 AGAGAGATTAAATGATGATGATTCATTAATTTCAAGATCATTAAATGATAGGGGAATAT 176
Db 104809 ACATAATTTTAAATTTAGAAATTTAGCAAAATCGTTCAATCCCTTTTGGAAAAA 104868
QY 177 CAAGCAATAATAAGCAATAAAGGCAATTAAGTAATAAGGAGAGATGATTCACCAATAT 236
Db 104869 CGATCTTAACCACTAGTAAATGTTAAGTCCTGCAATGACACGGGTTAATAAATTT 104928
QY 237 TGAATGAGGTGATGATTTCTTTTGCATGATGATGATGCGCAAACTACTAGATG 296
Db 104929 AGGATGACATGGCTCAGGCTTTTAAAGACAAATGACACTGATAGCTGTTCATAAAC 104988
QY 297 TTGGGACCCCTTCTCGATCTAATGAAAAAAGTATGGAATAGTAGATA 343
Db 104989 TTGTTATATTGCTTGAGAAATGAGGCCCATATATATTGATATATA 105035
```

```
RESULT 13
US-09-410-464-1
; Sequence 1, Application US/09410464
; Patent No. 635892
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: Poplar and other plant species.
```

FILE REFERENCE: 53375
CURRENT APPLICATION NUMBER: US/09/410,464
CURRENT FILING DATE: 1999-10-01
EARLIER APPLICATION NUMBER: 09/287,700
EARLIER FILING DATE: 1998-04-06
EARLIER APPLICATION NUMBER: 60/080,851
EARLIER FILING DATE: 1998-04-06
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1
LENGTH: 4285
TYPE: DNA
ORGANISM: Populus balsamifera subsp. trichocarpa
US-09-410-464-1

Query Match 4.4%; Score 43.8; DB 3; Length 4285;
Best Local Similarity 51.9%; Pred. No. 0.26;
Matches 122; Conservative 0; Mismatches 112; Indels 1; Gaps 1;
QY 684 ATACAATAATGACAAAATGCTTCCAAAGAA-AAACATGGCTCTTATAGTGAATATCG 742
DB 1378 ATCAAGGATAATTCAAAATTCGATCCAAATAACACATAGTAATGGAAGGACTTATGG 1437
QY 743 TTGACTGTTATAGAAAGATCTGAATTTATTATAAGAAATAGTGTCTTTCTTTCTTT 802
DB 1438 TATGAATGGATCAATAATATAAGGCTGAATTAACAACATTTTTTTTATTAGATCTGT 1497
QY 803 TCATATCTAAGGAGTAAGCAACCATGAATAGAAAAGGCTTAGTAATATATATCAAAAG 862
DB 1498 TTATTTTACGTTTAAAAATATTTTGAATTTATTTTATTTTATTTATATAAATTAAT 1557
QY 863 AATCGTCTTTTCTTTAAATATCGATAAAAATTTGTAATATAGATATAGAT 917
DB 1558 TTTTAGATCATTTTAAACGTAAATATAAAAATAATTTTTTTTAAAAAATTTAT 1612

RESULT 14
US-09-270-767-29271/c
Sequence 29271, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 29271
LENGTH: 392
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-29271

Query Match 4.4%; Score 43.6; DB 4; Length 392;
Best Local Similarity 50.5%; Pred. No. 0.15;
Matches 103; Conservative 1; Mismatches 100; Indels 0; Gaps 0;
QY 11 TTTGATTAAATTAATTAATTTGTATGACAAAGGATTAAACCTAGTTAATGATAATACT 70
DB 247 TATTATTACTCTACACCGGAGAATGAAAGCGAATAAACTAAATAGTCAGCGAAACG 188
QY 71 TCAGATCTATATCAATTAACAGCAATCACGGTCTATAGCAGCGTTGAGAGAAGATTAAAT 130
DB 187 TTACCAAGAATATCAGAAAGAGAAAGCAGTACACAAATGTAATACGACACATATGAT 128
QY 131 GTGATGTYCATTCATTAATTTCAAGATCATTATGATAGGGGATATCAAGCAATAATAA 190
DB 127 AAAATTACCATATAATTTTATGCAAGATAAATTTCTAGTTAAGAAGCAACAAAGTAAA 68
QY 191 CGATAAATGGCATTAAAGTAAATA 214
DB 67 TAATAAATAAGTTTGAAAAACA 44

RESULT 15

US-09-270-767-13327/c
Sequence 13327, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 13327
LENGTH: 2031
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-13327

Query Match 4.4%; Score 43.6; DB 4; Length 2031;
Best Local Similarity 50.5%; Pred. No. 0.24;
Matches 103; Conservative 1; Mismatches 100; Indels 0; Gaps 0;
QY 11 TTTGATTAAATTAATTAATTTGTATGACAAAGGATTAAACCTAGTTAATGATAATACT 70
DB 247 TATTATTACTCTACACCGGAGAATGAAAGCGAATAAACTAAATAGTCAGCGAAACG 188
QY 71 TCAGATCTATATCAATTAACAGCAATCACGGTCTATAGCAGCGTTGAGAGAAGATTAAAT 130
DB 187 TTAGCAAGAATATCAGAAAGAGAAAGCAGTACACAAATGTAATACGAAACACATATGAT 128
QY 131 GTGATGTYCATTCATTAATTTCAAGATCATTATGATAGGGGATATCAAGCAATAATAA 190
DB 127 AAAATTACCATATAATTTTATGCAAGATAAATTTCTAGTTAAGAAGCAACAAAGTAAA 68
QY 191 CGATAAATGGCATTAAAGTAAATA 214
DB 67 TAATAAATAAGTTTGAAAAACA 44

Search completed: November 11, 2004, 12:09:12
Job time : 109.152 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 10:27:15 ; Search time 535.761 Seconds
(without alignments)
10091.039 Million cell updates/sec

Title: US-10-009-966C-2_COPY_2155_3155
Perfect score: 1001
Sequence: 1 aatataagatttgattaaa.....accactactagaatccgg 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3625171 seqs, 2700493622 residues

Total number of hits satisfying chosen parameters: 7250342

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	55.6	5.6	158001	US-10-211-179-11	GENERAL INFORMATION
c 2	54.8	5.5	37515	US-10-433-793-28	Sequence 28, Appl
c 3	54.4	5.4	113515	US-10-311-455-2148	Sequence 2148, Ap
4	52.6	5.3	8201	US-10-240-485-64	Sequence 64, Appl
5	52	5.2	13584	US-10-311-455-588	Sequence 588, App
c 6	51.4	5.1	27890	US-10-741-601-5686	Sequence 5686, Ap
7	51.4	5.1	126872	US-10-741-601-5738	Sequence 5738, Ap
8	51.4	5.1	3673778	US-10-312-841-2	Sequence 2, Appli
9	50	5.0	5413	US-10-221-714A-418	Sequence 418, App
c 10	50	5.0	11745	US-10-240-453-206	Sequence 206, App
11	49.4	4.9	11694	US-10-221-714A-422	Sequence 422, App
c 12	49.4	4.9	3673778	US-10-312-841-2	Sequence 2, Appli

13	49.2	4.9	5360	15	US-10-204-708-66	Sequence 66, Appl
14	49.2	4.9	5360	15	US-10-311-455-1910	Sequence 1910, Ap
15	49.2	4.9	5360	15	US-10-240-589C-106	Sequence 106, App
c 16	49.2	4.9	5743	15	US-10-311-455-2041	Sequence 2041, Ap
17	49.2	4.9	7306	15	US-10-311-455-1609	Sequence 1609, Ap
c 18	49.2	4.9	34688	17	US-10-433-793-90	Sequence 90, Appl
19	49	4.9	9810	15	US-10-311-455-400	Sequence 400, App
20	48.8	4.9	6109	15	US-10-311-455-299	Sequence 299, App
c 21	48.8	4.9	6109	16	US-10-221-613-33	Sequence 33, Appl
c 22	48.8	4.9	34722	17	US-10-322-281-700	Sequence 700, App
c 23	48.4	4.8	12507	15	US-10-311-455-271	Sequence 271, App
c 24	48.4	4.8	37515	17	US-10-433-793-27	Sequence 27, Appl
25	48.2	4.8	7461	15	US-10-311-455-1757	Sequence 1757, Ap
26	47.8	4.8	15732	14	US-10-238-676-95	Sequence 95, Appl
27	47.8	4.8	15732	15	US-10-240-453-107	Sequence 107, App
c 28	47.8	4.8	48509	17	US-10-322-281-510	Sequence 510, App
29	47.8	4.8	3673778	15	US-10-312-841-1	Sequence 1, Appli
30	47.2	4.7	1223197	13	US-10-027-632-179264	Sequence 179264,
31	47.2	4.7	1223197	15	US-10-027-632-179264	Sequence 179264,
32	47	4.7	5572	15	US-10-311-455-1397	Sequence 1397, Ap
33	47	4.7	5572	15	US-10-240-452-59	Sequence 59, Appl
34	47	4.7	5979	14	US-10-239-676-18	Sequence 18, Appl
35	47	4.7	5979	15	US-10-240-453-28	Sequence 28, Appl
36	47	4.7	11422	15	US-10-311-455-192	Sequence 192, App
37	47	4.7	11422	16	US-10-257-166-18	Sequence 18, Appl
38	46.8	4.7	14316	16	US-10-221-613-408	Sequence 408, App
c 39	46.6	4.7	40862	15	US-10-311-455-2046	Sequence 2046, Ap
40	46	4.6	13328	15	US-10-311-455-1685	Sequence 1685, Ap
41	46	4.6	20579	17	US-10-433-793-104	Sequence 104, App
c 42	45.8	4.6	5184	16	US-10-221-613-383	Sequence 383, App
c 43	45.8	4.6	16545	15	US-10-311-455-23	Sequence 23, Appl
c 44	45.8	4.6	122614	13	US-10-087-192-1726	Sequence 1726, Ap
45	45.6	4.6	6240	15	US-10-311-455-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1

US-10-211-179-11

- ; GENERAL INFORMATION:
- ; APPLICANT: Nicholas M. Dean
- ; APPLICANT: Kenneth W. Dobie
- ; FILE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOTRYSYL PHOSPHATASE ACTIVATOR EXPRI
- ; TITLE REFERENCE: PTS-0011
- ; CURRENT APPLICATION NUMBER: US/10/211,179
- ; CURRENT FILING DATE: 2002-08-01
- ; NUMBER OF SEQ ID NOS: 119
- ; OTHER INFORMATION: n = a, t, c, or g
- US-10-211-179-11

Query Match 5.6%; Score 55.6; DB 16; Length 158001;

Best Local Similarity 42.6%; Pred No. 0.67;

Matches 346; Conservative 1; Mismatches 460; Indels 5; Gaps 1;

QY	117	AGAGAAGATTAAATGGATGTCATTCATTAATTTCAAGATCATTAAATGATGCGGAATAT	176
Db	22124	ATATATAATAATAA	22183

QY	177	CAAGCAATAAATCAAGTAAATGGCAATTAAGTAAATGAAGGAGATGATCCACCAATAT	236
Db	22184	TAATATATAATAA	22243

QY	237	TGAATGAGGTGATGATTCCTCTTTTGCATGATGATGATGCGCAATAGTACTAGATG	296
Db	22244	ATAATATAAAATAATATATAAATAATATAAATAATATAATATATATATATATATATATAA	22303

QY	297	TTGGGACCCCTTCGGAICTAATGAAAAAGATGGAATAGTAGTAATCGAATCTCTTT	356
Db	22304	ATAAAAATA	22363

QY	357	AGAAAGGTAGTGTCTCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGATATTTTFA	416
----	-----	--	-----

RESULT 10
US-10-240-453-206/c
; Sequence 206, Application US/10240453
; Publication No. US20030148326A1

```

? CURRENT FILING DATE: 2002-10-02
? PRIOR APPLICATION NUMBER: PCT/EP01/039973
? PRIOR FILING DATE: 2001-04-06
? PRIOR APPLICATION NUMBER: DE 10019058.8
? PRIOR FILING DATE: 2000-04-06
? PRIOR APPLICATION NUMBER: DE 10019173.8
? PRIOR FILING DATE: 2000-04-07
? PRIOR APPLICATION NUMBER: DE 10032529.7
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: DE 10043826.1
? PRIOR FILING DATE: 2000-09-01
? NUMBER OF SEQ ID NOS: 350
? SEQ ID NO 206
? LENGTH: 11745

```

Qy	13	TGATTAAATTAAATTAATATTGTATGACAAAGGATTAAAGCTAGTTAATGATATAACTTC	72
Db	6199	TTAACATATCTTTAAAAATCAATTAATTAATAAANAATCGAAAAACCAATATAAAAAAAA	6140
Qy	73	AGATCTATATCAAAATTAAACAGCAATCACGGTCATAGCAGC-GTTTGAGAGAAGATTAAATG	131

6139	Db	ATATATACTATTAAATTTTAAAAAATCAATATTAAATAATCTCTTAAAAAATAATAATAAAT	608
132	Qy	TGATGTCATTCATTAATTTCAAGATCATTTAAATGATAGGGGAATATCAAGCAATAAATAAC	151
6079	Db	TATATATCATCTTAAAAAATAATTTAAAAATATATACAATCAATTAATTTAAATAAT	6020
192	Qy	GATAAATGGCATTTAAAGTAAAGGAGAGAAATGATTCACCCAATATTGAATGAGGTGGATG	251
6019	Db	AUCTACAAATATACTATACITTTTAAATAATTAATAATAATTAATAATAATTAATAATAA	5960
252	Qy	ATCTCTCTTTTTGACAATCATGAATGATGGGCAAAATCATAGAAATGTTGGGACCCCTTCTCG	311
5959	Db	AACTCCCTCTAATTCCTATAACTAAAAATCTAAAAAAATCAAAAAATTTTAAAAAAATTAAC	5900
312	Qy	GATCTAATGA--AAAAAGTATGAATAGTAGATAATTCGAATCTCTTTAGAAAGGTAGTGA	369
5899	Db	CATTTTAATTAATTAATAATTTCTTTTAATTAACAAATAATTAATAAAATCTATAAAAAATTAATAAT	5840

QY 370 TTGCTTTTATCTAGAGAAAGTCTGCTTTTCAAGAAATATTTTATCAGAGAAATATTA 429
 Db 5839 TATCTATAAAAAATATATATAAAAACTTACAAAAATCCATAATTAATAAAAAATTTATA 5780
 QY 430 CA 431
 Db 5779 AA 5778

RESULT 11

US-10-221-714A-422
 ; Sequence 422, Application US/10221714A
 ; Publication No. US20040048254A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with
 ; TITLE OF INVENTION: tumor suppressor genes and oncogenes
 ; FILE REFERENCE: 5013.1005
 ; CURRENT FILING DATE: 2003-01-21
 ; CURRENT FILING DATE: 2003-01-21
 ; PRIOR APPLICATION NUMBER: PCT/EP01/02955
 ; PRIOR FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: DE 10013847.0
 ; PRIOR FILING DATE: 2000-03-15
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 540
 ; SEQ ID NO 422
 ; LENGTH: 11694
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-221-714A-422

Query Match 4.9%; Score 49.4; DB 16; Length 11694;
 Best Local Similarity 51.3%; Pred. No. 4.8;
 Matches 139; Conservative 0; Mismatches 131; Indels 1; Gaps 1;
 QY 12 TTGATTAATAATTAATTAATTTGATGACAAAGGATTAACCTAGT-TAATGATTAATACT 70
 Db 9647 TTGCGTAAAGATTAAATTTAATTAATAGGAAATTAATTAATTAATTTACG 9706
 QY 71 TCAGATCTATATCAATTAACAGCAATCACCGTCATAGCAGTGTGAGAGAAATTAAT 130
 Db 9707 TGATTAATAAAGTAGATTTATATATATAGAGTGGGAAAGATATTTAAGATAATAAAA 9766
 QY 131 GTGATGTCATTCATATTTCAAGATCATTAATGATGGGGAATATCAAGCAATAATAA 190
 Db 9767 ATGAATAAGTTTGAATTTTAAATTAATTTAATAAGAGGAAATTTAGAAATAATAA 9826
 QY 191 CGATAAAATGGCAATTAAGTAATAAGGAGATGATTCACCAATTAATTAATTAATTAAT 250
 Db 9827 ATGGATGGGAATTTAGAAATTAAGTCTTAATAGTAATTTATTTAGTAGGATTGTGA 9886
 QY 251 GATTCCTCTTTTGACAATGATGATG 281
 Db 9887 GGTTCCTTTTATTTATTTATTTT 9917

RESULT 12

US-10-312-841-2/c
 ; Sequence 2, Application US/10312841
 ; Publication No. US20030186277A1

; GENERAL INFORMATION:
 ; APPLICANT: Epigenomics AG
 ; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
 ; FILE REFERENCE: E01/1208/WO
 ; CURRENT APPLICATION NUMBER: US/10/312,841
 ; CURRENT FILING DATE: 2002-12-30
 ; NUMBER OF SEQ ID NOS: 2
 ; SEQ ID NO 2
 ; LENGTH: 3673778
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (379615)
 US-10-312-841-2

Query Match 4.9%; Score 49.4; DB 15; Length 3673778;
 Best Local Similarity 47.3%; Pred. No. 50;
 Matches 181; Conservative 0; Mismatches 201; Indels 1; Gaps 1;

QY 140 ATTCAATATTTCAAGATCATTAATGATAGGGAATATCAAGCAATTAATAACGATAATG 199
 Db 574390 ATTTAATTTTATATCTATCTCAAAATCCAAATATAAAAAATAAAAAATAAAAA 574331
 QY 200 GCATTAAAGTAATTAAGGAGATGATTCACCCATATTTGAATGAGGTGATGATCTTCT 259
 Db 574330 AAAATATCAATAAATAATAATTTAAAAAATTTCCAAATTTAAACAAAAACATAA 574271
 QY 260 TTTTGACATGATGATGATGGCAATATCTAGAAATGTTGGACCTTCTCGGATCTAAT 319
 Db 574270 AATTTAAAAATAATAATAAACCCCAATAAACAACCAACCAACCAACCAATC 574211
 QY 320 GAAAAAGTATGGAATAGTAGTAATTCGAATCTCTTTAGAAAGTAGTGTGCTTTTA 379
 Db 574210 ATAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAACA 574151
 QY 380 TCTAGAGAGAAATGCTGCTTTTCAAGATATTTTATCAGAGATATTTACATCCCTC 439
 Db 574150 TCTAACAAATAAAAAATAAACAACCAAAATTTTCATCAAAACTATATAAACCCGAA 574091
 QY 440 TCTCCCTATCTCTTTTCTTATTTATATGGACATTCCTCAATCAATCTTAAAGTACATA 499
 Db 574090 AAAATCACCACCTTTTCCAAATATAAAAAATAAATAAATAAATAAATAAATAA 574031
 QY 500 -CACCAAGATATTCATAAAT 521
 Db 574030 TCACCAAAAAATATCCAAAAAAT 574008

RESULT 13

US-10-204-708-66
 ; Sequence 66, Application US/10204708
 ; Publication No. US20030141852A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT FILING DATE: 2003-05-06
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1

```
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 66
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-66

Query Match          4.9%; Score 49.2; DB 15; Length 5360;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 108; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 684 ATACAATAATATGACAAAATTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAATAATTTGGATATAGTAGTATGGTTATTATTATTAAGATGATA 1852

QY 744 TAGACTCTTATAGAAAGATCTGAATTTATTTAAGAATAGTGTTCCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAGTTTTATGTAAGATATATTTTATTATTAGTTTA 1912

QY 804 CATATCTAAGGAGTAAGCAACCATCAATAGAAAGGCTTAGTAATATATATCAAGGA 863
Db 1973 AGTTGTTTATTGTAATTTGTAATTTGGAAT 1998

RESULT 14
US-10-311-455-1910
; Sequence 1910, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting Cytosine Methylation
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1910
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1910

Query Match          4.9%; Score 49.2; DB 15; Length 5360;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 108; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 684 ATACAATAATATGACAAAATTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAATAATTTGGATATAGTAGTATGGTTATTATTATTAAGATGATA 1852

QY 744 TAGACTCTTATAGAAAGATCTGAATTTATTTAAGAATAGTGTTCCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAGTTTTATGTAAGATATATTTTATTATTAGTTTA 1912

QY 804 CATATCTAAGGAGTAAGCAACCATCAATAGAAAGGCTTAGTAATATATATCAAGGA 863
Db 1973 AGTTGTTTATTGTAATTTGTAATTTGGAAT 1998

RESULT 15
US-10-240-589C-106
; Sequence 106, Application US/10240589C
; Publication No. US20040076956A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Repair
; TITLE OF INVENTION: DNA repair
; FILE REFERENCE: 5013.1008
; CURRENT APPLICATION NUMBER: US/10/240.589C
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03972
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 148
; SEQ ID NO 106
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-589C-106

Query Match          4.9%; Score 49.2; DB 16; Length 5360;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 108; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 684 ATACAATAATATGACAAAATTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAATAATTTGGATATAGTAGTATGGTTATTATTATTAAGATGATA 1852

QY 744 TAGACTCTTATAGAAAGATCTGAATTTATTTAAGAATAGTGTTCCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAGTTTTATGTAAGATATATTTTATTATTAGTTTA 1912

QY 804 CATATCTAAGGAGTAAGCAACCATCAATAGAAAGGCTTAGTAATATATATCAAGGA 863
Db 1913 GTTTTAAACGTTAGGAAAGAGTATATATTAAGAAAGTATTTTATGGTAATATTA 1972

QY 864 ATGGTGTTCCTTTTAAATATGAT 889
Db 1973 AGTTGTTTATTGTAATTTGTAATTTGGAAT 1998

Search completed: November 11, 2004, 14:31:46
Job time : 548.761 secs
```

This Page Blank (uspto)